



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 126551

TO: Phuong Bui
Location: REM-2A15/3C18
Art Unit: 1638
Friday, July 09, 2004

Case Serial Number: 09/857525

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Bui,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 01:29:12 ; Search time 6278 Seconds
(without alignments)
10880.634 Million cell updates/sec

Title: US-09-857-525C-1

Perfect score: 1576

Sequence: 1 gcacgagggaacatttgcg.....tcaaaaaaaaaaaaaaaa 1576

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.higo.hum.*
40: em.higo.mus.*
41: em.higo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1504.6	95.5	2156	8	AF276086	AF276086 Zea mays
2	1234	78.3	2247	8	AF276085	AF276085 Zea mays
3	987.4	62.7	2145	8	AK102647	AK102647 Oryza sat
4	987.4	62.7	2565	8	AK111501	AK111501 Oryza sat
5	979.4	62.1	1632	8	BT009111	BT009111 Triticum
6	577.8	36.7	1464	8	BT000625	BT000625 Arabidops
7	577.8	36.7	2090	8	AF250335	AF250335 Arabidops
8	577.8	36.7	2301	8	AF439826	AF439826 Arabidops
10	525.8	33.4	1887	8	AY247268	AY247268 Medicago
11	241.6	15.3	1923	8	AK121122	AK121122 Triticum
12	135	8.6	92392	8	AC128647	AC128647 Oryza sat
13	135	8.6	145115	8	AC096690	AC096690 Oryza sat
14	87	5.5	107857	2	AC146552	AC146552 Medicago
15	77.2	4.9	110514	8	F7G19	AC000106 Sequence
16	73.6	4.7	139089	2	AC146571	AC146571 Medicago
17	55.8	3.5	90373	3	AC115579	AC115579 Dictyoste
18	52.2	3.3	945	11	CNS06ICE	AL400116 T7 end of
19	50.6	3.2	1905	6	AX482695	AX482695 Sequence
20	50.6	3.2	2082	3	AF094764	AF094764 Drosophil
21	50.6	3.2	2577	3	BT001510	BT001510 Drosophil
22	50.6	3.2	2684	3	AY166755	AY166755 Drosophil
23	50.6	3.2	2687	3	BT003249	BT003249 Drosophil
24	50.6	3.2	2871	3	AY166754	AY166754 Drosophil
25	50.6	3.2	3157	3	AY166758	AY166758 Drosophil
26	50.6	3.2	3199	3	AY166753	AY166753 Drosophil
27	50.6	3.2	3210	3	AY084138	AY084138 Drosophil
28	50.6	3.2	3318	3	AY166757	AY166757 Drosophil
29	50.6	3.2	3497	3	AF094763	AF094763 Drosophil
30	50.6	3.2	3538	3	BT010233	BT010233 Drosophil
31	50.6	3.2	3679	3	AY166756	AY166756 Drosophil
32	50.6	3.2	4387	3	BT003626	BT003626 Drosophil
33	50.6	3.2	4666	3	AY070541	AY070541 Drosophil
34	50.6	3.2	4783	3	AY166752	AY166752 Drosophil
35	50.6	3.2	80069	2	AC019671	AC019671 Drosophil
36	50.6	3.2	173634	3	AC009344	AC009344 Drosophil
37	50.6	3.2	195868	3	AC008308	AC008308 Drosophil
38	50.6	3.2	235707	3	AE003733	AE003733 Drosophil
39	49.6	3.1	969	6	AX595096	AX595096 Sequence
40	49.6	3.1	969	6	AX818518	AX818518 Sequence
41	49.6	3.1	969	6	AX829548	AX829548 Sequence
42	49.6	3.1	1434	8	YSCSNF4	M30470 Yeast (S.ce
43	49.6	3.1	1619	8	YSCCAT3	M21760 Yeast (S.ce
44	49.6	3.1	1818	8	SCYGL115W	W27637 S.cerevisia
45	49.6	3.1	163443	2	AC006280	AC006280 Plasmodiu

ALIGNMENTS

RESULT 1
AF276086
LOCUS
DEFINITION
AF276086
ACCESSION
AF276086.1
VERSION
AF276086.1
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2156)
Lumbreras, V., Alba, M.M., Kleinow, T., Koncz, C. and Pages, M.
Domain fusion between SNF1-related kinase subunits during plant

evolution
 JOURNAL EMBL Rep. 2 (1), 55-60 (2001)
 MEDLINE 21146516
 PUBMED 11252725
 REFERENCE 2 (bases 1 to 2156)
 AUTHORS Lumbres, V. and Pages, M.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2000) Genetica Molecular, CID, CSIC, Jordi Girona,
 18-26, Barcelona 08034, Spain
 FEATURES
 Location/Qualifiers
 1..2156
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 /dev_stage="seedling"
 /note="under water stress"
 456..1946
 /note="Scf-4 like"
 /codon_start=1
 /product="protein kinase AKINbeta gamma-2"
 /protein_id="AAG31752.1"
 /db_xref="GI:11139548"
 /translation="MPSHGADSAHDAGAVGVSSGATVPTFVWPYGGKRVFVSGSPT
 RWSEHLPMSIEGPTFQAIICSLSPGHEIKFVDGEMRDERQPTISGFGIVNTL
 YLREYNQIMTLESPSPGSMNDVNEFQRTVTLSDGTVEGTLRVSEAAIQISR
 CRSEYLSLNTCYDLLPDSKVIADINLPVKOSFHLHSGQIPVAPLWSPFRGQV
 LLSPLDFILIRLETHGSLNTEQLETHLSWKEAKROTNRGRNDQWRPQOHLVHA
 TPYERLDIAVLKQNGISIVPIYSSSDGSPFOLLHLASLSGLKLCICRYFKNSTG
 NLPILOPVCISIPGSPVPGKIDNLSPLMLPNSLSALNMLVQAGVSSPIVDD
 NDSLDYTSRSDITAKAKVYTHRLDENTHQAQLQGDATPFQFFNGRCQMCL
 RSDPLKLMERLANPGVRVFEIVAGSKRVEGIIISLIDIEKFLLSL"

CDS

ORIGIN

Query Match 95.5%; Score 1504.6; DB 8; Length 2156;
 Best local Similarity 99.0%; Pred. No. 0;
 Matches 1537; Conservative 2; Mismatches 24; Indels 8; Gaps 1;
 QY 8 GGAACATTTCCGATGCTCTCTATCGAAGGCTGCCACCTGTATTTTCAGGCTATTTTCAG 67
 DB 596 GGAACATTTCCGATGCTCTCTATCGAAGGCTGCCACCTGTATTTTCAGGCTATTTTCAG 655
 QY 68 TCTGTCTCTCGGATTCACGAGTACAAATTTCTTTGTGGACGCGGAAATGGCGCATGATGA 127
 DB 656 TCTGTCTCTCGGATTCACGAGTACAAATTTCTTTGTGGACGCGGAAATGGCGCATGATGA 715
 QY 128 GCGTCAACCTACATCTCTCGGAGTTTGGCATAGTTTAAACACTTTACTTTGACAAGGGA 187
 DB 716 GCGTCAACCTACCATATCTCGGAGTTTGGCATAGTTTAAACACTTTACTTTGACAAGGGA 775
 QY 188 ATATAACCAATAAACCCTTATCAAGTCCAAAGCACACCTGGAGCAGGATGAACATGGA 247
 DB 776 ATATAACCAATAAACCCTTATCAAGTCCAAAGCACACCTGGAGCAGGATGAACATGGA 835
 QY 248 TGTGGATAATGAAATTTTCAACGTACGTTTACGTTCTCAGATGGCCACCGTTTCAGAAAG 307
 DB 836 TGTGGATAATGAAATTTTCAACGTACGTTTACGTTCTCAGATGGCCACCGTTTCAGAAAG 895
 QY 308 TACTCTGAGAGTTTCAGAGCTGCAATACAAATATCTAGGTCGCTGTTCTGATATCT 367
 DB 896 TACTCTGAGAGTTTCAGAGCTGCAATACAAATATCTAGGTCGCTGTTCTGATATCT 955
 QY 368 GAATTTGCATACATGCTATGATTTACTCCAGATTCTGGCAAGGTTATTCGCCCTAGACAT 427
 DB 956 GAGTTTGCATACATGCTATGATTTACTCCAGATTCTGGCAAGGTTATTCGCCCTAGACAT 1015
 QY 428 TAATTTACCTGTGAAGCAATTCATTCATATTCCTCCATGAACAGGGGATTCCTGAGCTCC 487
 DB 1016 TAATTTACCTGTGAAGCAATTCATTCATATTCCTCCATGAACAGGGGATTCCTGAGCTCC 1075
 QY 488 TCTCTGAGACTCATTCAGAGGTCATATTTGTTGCTCTCTAGCCCATTTGGATTTTCATCT 547
 DB 1076 TCTCTGAGACTCATTCAGAGGTCATATTTGTTGCTCTCTAGCCCATTTGGATTTTCATCT 1135

QY 548 CATATTGCGGGAGCTAGAACTCATGCTGGAACCTTGACAGAGCAGCTTGAACACA 607
 DB 1136 CATATTGCGGGAGCTAGAACTCATGCTGGAACCTTGACAGAGCAGCTTGAACACA 1195
 QY 608 CACTATATCTGCATGSAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATGTCAGTG 667
 DB 1196 CACTATATCTGCATGSAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATGTCAGTG 1255
 QY 668 GCGACCGCAACAGCATCTAGTGCATGCCACCCCTTATGAGTCCTTGAGGGAATGTCAGT 727
 DB 1256 GCGACCGCAACAGCATCTAGTGCATGCCACCCCTTATGAGTCCTTGAGGGAATGTCAGT 1315
 QY 728 AAGCTTTTGCAGAAATGSCATTTCTACAGTGCACAGTTATTTATTCATCATCATCAGATGG 787
 DB 1316 AAGCTTTTGCAGAAATGSCATTTCTACAGTGCACAGTTATTTATTCATCATCATCAGATGG 1375
 QY 788 ATCATTTCCGAGTTATTGTCATCTTGCATCACTTTCTGGAAATTTTGAATGATTTTGTAG 847
 DB 1376 ATCATTTCCGAGTTATTGTCATCTTGCATCACTTTCTGGAAATTTTGAATGATTTTGTAG 1435
 QY 848 ATACTTCAAAAACCTCAACTGTAATTTGCTGCTTGTGAACCAACAGTGTGCTCCATTC 907
 DB 1436 ATACTTCAAAAACCTCAACTGTAATTTGCTGCTTGTGAACCAACAGTGTGCTCCATTC 1495
 QY 908 GCTGGGTTCTCTGGGTTCCGAAAATTTGGTGATCTCAACAGTCTGTCATTGGGCTATGTTGCG 967
 DB 1496 GCTGGGTTCTCTGGGTTCCGAAAATTTGGTGATCTCAACAGTCTGTCATTGGGCTATGTTGCG 1555
 QY 968 ACCTAATGCTCACTTAGCTCTGCTTAAACATGTTGGTTCCTCAAGCTGGAGTAAGCTCAAT 1027
 DB 1556 ACCTAATGCTCACTTAGCTCTGCTTAAACATGTTGGTTCCTCAAGCTGGAGTAAGCTCAAT 1615
 QY 1028 ACCAATTTGTCGATGACAAACGCTCCCTGCTTGACACTTCTCTAGAGTGCATCAGAC 1087
 DB 1616 ACCAATTTGTCGATGACAAACGCTCCCTGCTTGACACTTCTCTAGAGTGCATCAGAC 1675
 QY 1088 CCTAGCAAAAAGACAAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACCGAGC 1147
 DB 1676 CCTAGCAAAAAGACAAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACCGAGC 1735
 QY 1148 TTTGACGCTTTGGAACAAGTCCCAATACACCTTTTGGATTTCTTTAAACGCGCAGAGATGCCA 1207
 DB 1736 TTTGACGCTTTGGAACAAGTCCCAATACACCTTTTGGATTTCTTTAAACGCGCAGAGATGCCA 1795
 QY 1208 GATGTCCTCGGCTGATCTCTTCTGAGAGTGTGAGGCGCATGCGCTAATCTCGGGT 1267
 DB 1796 GATGTCCTCGGCTGATCTCTTCTGAGAGTGTGAGGCGCATGCGCTAATCTCGGGT 1855
 QY 1268 GCGCGGGTGTTCATTTGTGGAAGCTGGGAGCAAACTGTGAGGCGCATCATCATCAAG 1327
 DB 1856 GCGCGGGTGTTCATTTGTGGAAGCTGGGAGCAAACTGTGAGGCGCATCATCATCAAG 1915
 QY 1328 TGACATTTTCAAGTTCTGTCGAGCTTGTGAGGAGAGAGGATGCGGAGTTTTCCTTTT 1387
 DB 1916 TGATATTTTCAAGTTCTTGTGAGCTTGTGAGGAGAGGAGCATGCGGAGTTTTCCTTTT 1975
 QY 1388 AACAGTAGCAACCTCTGCGGCTTTGGGCTCCTTGCGATTAATTCGCTATCTGTCAGGACT 1447
 DB 1976 AACAGTAGCAACCTCTGCGGCTTTGGGCTCCTTGCGATTAATTCGCTATCTGTCAGGACT 2027
 QY 1448 CAGGTTTCATGCTGTGAAAGTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTT 1507
 DB 2028 CAGGTTTCATGCTGTGAAAGTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTT 2087
 QY 1508 CCGTTTGTAAACAAACATACATCAGGTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTT 1567
 DB 2088 CCGTTTGTAAACAAACATACATCAGGTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTT 2147
 QY 1568 AAAAAAAA 1576
 DB 2148 AAAAAAAA 2156

```

RESULT 2
LOCUS AF276085 2247 bp mRNA linear PM 13-AUG-2002
DEFINITION Zea mays protein kinase AKINbetagamma-1 mRNA, complete cds.
ACCESSION AF276085
VERSION AF276085.1 GI:11139545
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2247)
AUTHORS Lumbieras V., Alba, X.M., Kleinow, T., Koncz, C. and Pages, M.
TITLE Domain fusion between SNF1-related kinase subunits during plant
evolution
JOURNAL EMBO Rep. 2 (1), 55-60 (2001)
MEDLINE 21146516
PUBMED 11252725
REFERENCE 2 (bases 1 to 2247)
AUTHORS Lumbieras, V. and Pages, M.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2000) Genetica Molecular, CIB, CSIC, Jordi Girona,
18-26, Barcelona 08034, Spain
FEATURES
source
1..2247
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/dev_stage="seedling"
/note="under water stress"
496..1989
/feature="Snf-4 like"
/codon_start=1
/product="protein kinase AKINbetagamma-1"
/protein_id="AAAG31751.1"
/db_xref="GI:11139545"
/translation="MFSGADSDAGDTGVSSGVGATVPARFVWPYGGKRVFVSGSF
TWRSHELPMSVEGCTPQALCSLSPGHEVYKVVDCGEMHDEROPTISGRFGIVNT
LYLREYFQINLALNPSTPGSRMMNDVNFCHVTLSLDGHIPEGVKVSMAAIOIS
RCEVSEYLNHLDPDPSGKVIADLNLPVKQSFHILHEOGI PVAPWDSFRQGFV
GLSLPLDLILRLREIETHGNSITDQLEHTHSAMKEAKROTCEGRNDGWRHRLDVLH
ATPVESLRDIAVKKLLNDISTVPIYSSSSDGSFPOLLHLASLGLKICIFRYFKNST
GNLPLNPGVSGIPGSPVKIPDPSPLMLRPNASLSALNMLVOAGVSSIPITVD
ENDSLDYSRSDITLAKAKUYTHVRLEMAIHOALQLOQDANTPEGFNQRQCMC
LRSDPLKMERLANPGVRRVFIENGSKRVGGIISLSDIFKFLSL"
ORIGIN
Query Match 78.3%; Score 1234; DB 8; Length 2247;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 155; Indels 37; Gaps 5;
QY 8 GGAACTTGGCGATGTCCTCATCGAAGGCTGCCCACTGTATTTCAGCTATTGTCAG 67
DB 639 GGAACTTGGCGATGTCCTCATCGAAGGCTGCCCACTGTATTTCAGCTATTGTCAG 698
QY 68 TCTGTCTCTCGGATTCAAGATACAAATCTTTGTGGACGGGATGGCGCATGATGA 127
DB 699 CTTGTCTCAGGATTCATGATACAAATCTTTGTGGACGGGATGGCGCATGATGA 758
QY 128 GCGTCAACCTACCATATCTGGGAGTTTGGCATAGTTAAACACATTTACTTGCAGAGGA 187
DB 759 GCGCAACCTACTATATCTGGGAGTTTGGCATAGTTAAACACATTTACTTGCAGAGGA 818
QY 188 ATATACCAATTAACACCTTATCAAGTCCAGACACACCTGGACGAGATGACATGGA 247
DB 819 ATTTAATCAATTAATGCTTTTAAATCCAGACACACCTGGACGAGATGACATGGA 878
QY 248 TGTGATTAATGAATAATTTTCAACAGTACGCTTACGTTGTGATGTCAGATGGCAGCGTTTCAAGG 307
DB 879 TGTGATTAATGAATAATTTTCAACAGTACGCTTACGTTGTGATGTCAGATGGCAGCGTTTCAAGG 938
QY 308 TACTCTGAGAGTTTCAGAGGCTGCAATACAAAATATCTAGGTGTCTGTGTTTCTGCAATATCT 367

```

```

939 TACTCTGAGAGTTTCAGAGGCTGCAATACAAAATCTCAAGTGCCTGTTTCTGAGTATCT 998
QY 368 GAATTTGCATACATGCTATGATTTACTCCAGATTTCTGGCAAGGTTATTCCTTACACAT 427
DB 999 CAATTTGCATACATGCTATGATTTACTCCGGAATTCGGGCAAGGTTATTCCTTGCACAT 1058
QY 428 TAATTTACCTGTGAAGCAATTCATTCATATTCCTCAATGAACGGGATTCCTTGTAGTCTC 487
DB 1059 TAATTTACTCTGTGAAGCAATTCCTTCTATTCATGATGAACGGGATTCCTTGTAGTCTC 1118
QY 488 TCTCTGGGACTCATTTACAGAGGTCATTTCTGTCTCTCTTCTAGCCCATTTGATTTCACT 547
DB 1119 TCTCTGGGACTCATTTACAGAGGTCATTTCTGTGGCCCTTCTGAGCCCATTTGATTTCACT 1178
QY 548 CATTTTGGGGAGCTAGAAAACCTCATGCTCGAATTCGACAGAGAGAGGATTCAGAACACA 607
DB 1179 CATTTTGGGGAGCTAGAAAACCTCATGCTCGAATTCGACAGAGAGATTCAGCTTGAACACA 1238
QY 608 CACTATATCTGCATGGAAGAGGCTTAAGCGGCAAACTAATGGAAGAAATGATAGTCAGTG 667
DB 1239 CACTATATCTGCATGGAAGAGGCTTAAGCGGCAAACTGTTGGAGAAATGATAGTCAGTG 1298
QY 668 GCGACCGCAACAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTGAGGACATTCAGT 727
DB 1299 GCGAGCACATCAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTGAGGACATTCAGT 1358
QY 728 AAAGCTTTTGCAAAATGGGATTTCTACAGTGGCAGTTATTTATTCATCATCATCAGATGG 787
DB 1359 AAAGCTTTTGTAAATGACATTTCTACAGTGGCAGTTATTTATTCATCATCATCAGATGG 1418
QY 788 ATCAATTCCTCGAGTATTGTCATCTTCATCATCTTCGAAATTTGAAATGATTTGTAG 847
DB 1419 ATCAATTCCTCGAGTATTGTCATCTTCATCATCTTCGAAATTTGAAATGATTTGTAG 1478
QY 848 ATACTTCAAAAACCTCAACTGGTAAATTTGCTATTTCTGAAACCAACCGATGTCCTCAATTC 907
DB 1479 GTATTTTAAAAACCTCAACTGGTAAATTTGCTATTTCTGAAACCAACCGATGTCCTCAATTC 1538
QY 908 GCTGGGTTCTGGGTTCCGAAAATTTGGTGTATCTGACAGTCTGCTCCATTTGCTATGTTGCG 967
DB 1539 GCTGGGTTCTGGGTTCCGAAAATTTGGTGTATCTGACAGTCTGCTCCATTTGCTATGTTGCG 1598
QY 968 ACCTAATGCTCACTTGTAGTCTGCTTAAACATGTTGTTTCAAGCTGGAGTAAAGTCAAT 1027
DB 1599 ACCTAATGCTCACTTGTAGTCTGCTTAAACATGTTGTTTCAAGCTGGAGTAAAGTCAAT 1658
QY 1028 ACCAATTTGGATGACACGATCTCCCTGCTTGTACATCTTCTAGAGTACATCATCAGC 1087
DB 1659 ACCAATTTGGATGACACGATCTCCCTGCTTGTACATCTTCTAGAGTACATCATCAGC 1718
QY 1088 CTTAGCAAGAACAAGGCTTACACATCTTGGCTGGATGAGATGACCAATTCACACAGGC 1147
DB 1719 CTTAGCTAAGCAAGGCTTACACATCTTGGCTGGATGAGATGAGATGAGATGAGATGAGATGAG 1778
QY 1148 TTTGAGCTTTGACAGATGCCAATACACCTTTTGGATTTTAAACGGGCGAGATGCCA 1207
DB 1779 TTTAGAGCTTTGACAGATGCCAATACACCTTTTGGATTTTAAACGGTCCAGATGCCA 1838
QY 1208 GATGTGCTCCGGTCTGATCTTGTGCTGAAGGATGAGGAGGAGTCTGCTAAATCTCGGGGT 1267
DB 1839 GATGTGCTCCGGTCTGATCTTGTGCTGAAGGATGAGGAGGAGTCTGCTAAATCTCGGGGT 1898
QY 1268 GCGGGGGTGTTCATTTGTGAAGCTGGGAGCAAACTGTGGAGGGGATCATATCACTAAG 1327
DB 1899 GCGGGGGTGTTCATTTGTGAAGCTGGGAGCAAACTGTGGAGGGGATTTATATCATCTAG 1358
QY 1328 TGACATTTTCAAGTCTTCTGCTGAGCTTTGAGGAG---AGAGGAGCATGGCGAGTTTTT 1384
DB 1959 TGATATTTTCAAGTCTTCTGCTGAGCTTTGATGAAGAGAGAGGAGCATGACGGACATTG 2018
QY 1385 TTT-----AACAGTAGCAACCTCTGGGGCTTTGGGCTCTTGGGATTAATTTGGCTATCG 1438

```


Db 560 TCTGAATGACAACTGGGTATGATTTACTCCAGATTCTGGCAGGTGCTTCTCTAGA 619
 QY 425 CATTATTTAATCTGTGAAGCAATCAATTCATTTCTCCATGAAACAGGGGATTCCTGTAGC 484
 Db 620 CTTTAAATTCCTGTGAAGCAATCTTTTCTATATCTTCTATGAAACAGGGGAAATTCCTGTGGC 679
 QY 485 TCCTCTCTGGGACTCATTCAGAGGTCAATTTGTGTCTCTCTAGCCCAATGGATTTCAT 544
 Db 680 ACCCTATGGGATTCATTGAGGGCCAGTTTGTGGCCCTTTGAGCCCACTGGATTTCAT 739
 QY 545 ACTCATATTTGCGGAGCTAGAACTCATGCTCGAACTTTGACAGAAGACGCTTGAAC 604
 Db 740 ACTCATATTTGAGAGAGCTGGAACCTCATGCTCCAACTCTGACAGAAGACGCTTGAAC 799
 QY 605 ACACACTATATCTGATGTAAGAGAGCTGAAGGGCAACCTAATGGAAGAAATGATGCTCA 664
 Db 800 ACATATATATCTGCTGGTGAAGGGCCCAAGCGTCAAACTTATGCGCAGAAACGAGGGTTC 859
 QY 665 GTGGGACCGCAACAGCATCTAGTCATGCGACCCCTTATGAGTCTCTGAGGACATTCG 724
 Db 860 TTGGAGGGCAATCACCATTATGCTCATGCCACCCCTTATGATCTCTGAGGGAATTCG 919
 QY 725 AGTAAAGCTTTTGCMAAATGGCATTTCTACAGTCCAGTATTTATTCATCATCATCAGA 784
 Db 920 TATGAAGTACTGCAAAACGGTCTTCTACCGTTCCAATATATGTTTCTCTCATCACCAGA 979
 QY 785 TGGATCATTTCCCGCAGTTATGCACTTCATCATCATCTTCTGGATTTTGAATGTTATTCG 844
 Db 980 TGGCTCATATCACAATTTGTCATCTGCTTCCCTTCTCGGAATTTTGAATGTTATTCG 1039
 QY 845 TAGATACCTTCAAAACCTCAACTGGTAAATTTGCTTATTTCTGAACCAACCACTGTCCTCCAT 904
 Db 1040 TAGATATTTTAAATTTCTCAAGTAAATTTACCTATTTTGAGCCCACTGTATGCAAT 1099
 QY 905 TCGGTGGGTCTCTGGGTTCCGAAATTTGTGTGATCTGAACAGTCGTCATTCATTCGCTAGCT 964
 Db 1100 TCTCTGGGTAGCTGGGTTCCAAATTTGTGTGATCTGAATGCGGCTCCATTCGCTATGCT 1159
 QY 965 GGAACCTAATGCTCATCTAGCTGCTGCTTACATGTTGGTTCAGCTGGAGTAGCTCC 1024
 Db 1160 GGGGCTTAACACATCTCTTAGCGCTGCTGCTGAAATTTGCTGGTTCAAGCTGGTGTAGCTC 1219
 QY 1025 AATACCAATTTGGATGACAAACGACTCCCTGCTGTGACACTTACTCTAGAGTGCATCATC 1084
 Db 1220 AATACCAATTTGGATGACAAAGACTCTGCTGTGACACATATCCAGAGTGCATCATC 1279
 QY 1085 AGCCCTAGCAAAAGCAAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACCA 1144
 Db 1280 AGCTCTTGTCAAAAGCAAGGTTGTACACACATTCGCTAGATGAGATGACCATTCATCA 1339
 QY 1145 GCGTTTGCAGCTTGGACAGATGCCAATACACCTTTTGGATCTTTAAAGCCGACAGATG 1204
 Db 1340 GGCCTGACGCTTGGACAGATGCCAATTCGCCCTTTGGATCTTTAAAGCCGACAGATG 1399
 QY 1205 CCAGATGTCCTCGCTGCTGATCTCTTGTGTGAGTGCATGAGCTGAGCTGATCTGCTG 1264
 Db 1400 TCAGATGTCCTCGCTGCTGATCTCTTGTGTGAGTGCATGAGCTGAGCTGATCTGCTG 1459
 QY 1265 GGTGGGCGGGTGTTCATTTGTGGAAGCTGAGCAACAGCTGTGGAGGGCATCTATCACT 1324
 Db 1460 GGTGGGCGGGTGTTCATTTGTGGAAGCTGAGCAACAGCTGTGGAGGGCATCTATCACT 1519
 QY 1325 AAGTGACATTTTCAAGTCTTCTGCTGAGCT 1353
 Db 1520 GAGTGTATTTTCAAGTCTTCTGCTGAGCT 1548

RESULT 4

AK111501

LOCUS

DEFINITION

Accession

AK111501

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLES

JOURNAL

COMMENT

AK111501.1 GI:37988164
 FLI CDNA: CAP trapper.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Euryotia; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

The Rice Full-length cDNA Consortium, National Institute of
 Agrobiological Sciences Rice Full-length cDNA Project Team,
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
 Ohtsuka, K., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group: Ohtsuka, Y., Murakami, K.,
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
 Narioka, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Yoshino, M., and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice
 Science 301 (5631), 376-379 (2003)

22752273

12869764

2

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
 Hori, F., Hotta, I., Iida, J., Ikeda, R., Ikehata, M., Inamura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y.,
 Nakamura, M., Namiki, T., Narioka, R., Niikura, J., Niishi, K.,
 Nomura, K., Numasaki, R., Ohneda, E., Ohtsuka, K., Oka, M.,
 Ooka, H., Osato, N., Ota, Y., Ohtsuka, K., Ohtsuka, K., Sakai, C.,
 Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,
 Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y.,
 Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M.,
 Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akashira, S.,
 Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K.,
 Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J.,
 Yokomizo, S., and Yoshimura, A.
 Rice full-length cDNA

Unpublished

3 (bases 1 to 2565)

Kikuchi, S.

Direct Submission

Submitted (12-SEP-2002) Shoshi Kikuchi, National Institute of
 Agrobiological Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
 Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 32K full-length cDNA clones from japonica
 rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T.,
 Yamamoto, M., and Nakahama, Y.
 FAIS Genome Sequencing & Analysis Group: Ohtsuka, Y., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M.,
 Kobayashi, M., Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M.,

AK111501 2565 bp mRNA linear PLN 29-OCT-2003

Oryza sativa (japonica cultivar-group) cDNA clone:J013000N04, full

insert sequence.

AK111501

Masuda, H., Miura, J., Mizuno, K., Nariawa, R., Niihara, J., Oka, M.,
 Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M.,
 Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Haragaki, T.,
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oosato, N.,
 Ota, Y., Saichoh, H., Sakai, C., Sakai, K., Sasaki, D., Sato, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M.,
 Tagami-Takeda, Y., Tagawa, A., Takahashi, E.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
 Yasunishi, A. and Hayashizaki, Y.

FEATURES

source
 1. 2565
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J01300N04"

ORIGIN

Query Match 52.7%; Score 987.4; DB 8; Length 2565;
 Best Local Similarity 83.2%; Pred. No. 3.4e-262;
 Matches 1123; Conservative 0; Mismatches 226; Indels 0; Gaps 0;
 5 GAGGAAATTTGGCGATGCTCTCTATCGAAGGTGCGCCACTGATTTTGGCTATTG 64
 610 GACAGAAATTTGGCAATGCTCTCTGTTGAAGGTTGTCCACCGTTTTCAGGGCAATG 669
 65 CAGTCTCTCTCGGATTCAGAGTACAAATCTTCTGTGACGGGAATGGCGCATGA 124
 670 CAGTTTGTCCCTGGGATTCATCAGTACAGTTTGTGTGATGGGAAATGGCGCAGA 729
 125 TGAGCGTCAACTACCATATCTGGGGAGTTGGCATAGTTTAAACACATTTTACTTGACAAG 184
 730 TGAGCGCACTACCATATACAGGAGCATAGTGTGTTAAACATTTTATGCTTGACTAG 789
 185 GGAATATAACCAATAACACTTATCAAGTCCAGGACACCTGGAACAGATGAACAT 244
 790 GGAATTTGACAGATAAATAAATATGAGCGCTAGTACACCTGGAAGTAGAGTAAACAT 849
 245 GGATGTGATATGAAATTTTCAACGTACGGTTACGTTGTGATGATGACACCGTTTCAGA 304
 850 GGATGTGACAAACGACAAATTTTCAAGTACGTTGTTCTTGTCTGATGATATTCAGGA 909
 305 AGGTACTCTGAGAGTTTCAGAGGTGCGCAATACAAATATCTAGGTGCTGTGTTCTGAATA 364
 910 AGGTCTCTCAGAGATTTTCAGAGGCGAGCTATACAGATCTCTAGGTGCTGTGAGCGGATT 969
 365 TCTGAATTTGATACATGCTGATTTACTTCCGATCTCGCAGGTTATTCGCTTGA 424
 970 TCTGAATGACAAATCTGGGTATGATTTACTTCCAGATCTCGCAAGGTCATGCTCTAGA 1029
 425 CATTAAATTTACTGTGAAGCAATCATTCATATTTCCATGAACAGGGGATTCCTGTAGC 484
 1030 CGTTAAATTTGCTGTGAAGCAATCTTTTATATCTTCTATGACAGGGAATTCCTGTGCG 1089
 485 TCTCTCTCGGACTCATTCAGAGGTCAATTTGTTGTTCTCTCTGACCAATTTGATTTTCA 544
 1090 ACTCTATGAGGATTCATTCAGGGGCGAGTTTGTGGCGCTTTTGGACCACTGGATTTTAT 1149
 545 ACTCATATTTGGGAGCTAGAACTCATGCTCGAATTTGACAGAGGAGCTTGAAC 604
 1150 ACTCATATTTGAGAGCTGGAACATCATGCTGCTTCAATCTGACAGAGGAGCTTGAAC 1209
 605 ACACACTATATCTGATGAAAGAGCTTAAGCGCAAACTAATGGAAGAAATGATGATCA 664

Db 1210 ACATCTATATCTGCTGGAGGAGGCGCAAGGTCACAACTTATGCCAGAAACGAGGGTTC 1269
 Qy 665 GTGCGGACCGCAACAGCATCTAGTGCATGCGACCCCTTATGAGTCCCTTGAGGACATGCG 724
 Db 1270 TTGAGGCGCAAAATCACCAATTTAGTTCAITGCGACCCCTTATGAATCCCTTGAGGAAATGCG 1329
 Qy 725 AGTAAAGCTTTTGCAGAAATGSCATTTCTACAGTGCAGTATTTATTTATCATCATCATCAGA 784
 Db 1330 TATGAGCATATCTGCAAAACGCTGTTCTTACCGTTCATATATGTTTCTCATCATCAGAGA 1389
 Qy 785 TGGATCATATCCCGCAGATTTATGATCTTTGATCATCTTTCTGGAATTTTGAATGATTTG 844
 Db 1390 TGGCTCATATCCCAATTTGTCATCTTCTTCTTCTGGAATTTTGAATGATTTG 1449
 Qy 845 TAGATCATCTCAAAATCTCAACTGCTGTAATTTGCTTATTTCTGAAACCAACAGTGTCTCCAT 904
 Db 1450 TAGATATTTTAAATAATTTCTCAAGGTAATTTACCTATTTTGGAGCAACCTGTATGCCAAT 1509
 Qy 905 TCCGCTGGGTTCTCGGTTCCGAAATTTGGTGAATCTGAACAGTCTGCTTCATTTGGCTATGTT 964
 Db 1510 TCCTCTGGGTACCTGGGTTCCAAATTTGGTGAATCTTAAATGCTGCTTCAATGCTATGTT 1569
 Qy 965 GCGACCTAATGCTCCTACCTAGTCTGCTTCTTAAATGCTGCTTCAATGCTGCTGCTGCT 1024
 Db 1570 GCGGCTTAAACATCTCTTAGCGCTGCTTGAATTTGCTGCTTCAATGCTGCTGCTGCT 1629
 Qy 1025 AATACCAATTTGCTGATGACAGTCCCTGCTTGAACATTTTCTAGAGTGAATGCTGCTGCT 1084
 Db 1630 AATACCAATTTGCTGATGACAGTCCCTGCTTGAACATTTTCTAGAGTGAATGCTGCTGCT 1689
 Qy 1085 AGCCTTAGCAAAAGCAAGGTTCTACACATGTTTGGCTGGATGAGATGACCATTCACCA 1144
 Db 1690 AGCTTTGCAAAAGCAAGGTTCTACACATGTTTGGCTGGATGAGATGACCATTCACCA 1749
 Qy 1145 GCGTTGCGCTTGGCAAGATGCAATATACCTTTTGGATTTCTTTAAGCGGACAGATG 1204
 Db 1750 GCGTCTGCGCTTGGCAAGATGCAATATACCTTTTGGATTTCTTTAAGCGGACAGATG 1809
 Qy 1205 CCAGATGCTCCGCTGCTGATCTTCTGCTGAAGTGTAGCGGACCTGGCTTAATCTCTG 1264
 Db 1810 TCAGATGCTCTCGGCTGCTGATCTTCTGCTGAAGTGTAGCGGACCTGGCTTAATCTCTG 1869
 Qy 1285 GGTGCGGCGGTGTTTCTTGTGGAAGTGTGGAGCAACAGTGTGGAGGCGATCATATCACT 1324
 Db 1870 GGTGCGGCGGTGTTTCTTGTGGAAGTGTGGAGCAACAGTGTGGAGGCGATCATATCACT 1929
 Qy 1325 AAGTGCATTTTCAAGTCTTCTGCTGAGCT 1353
 Db 1930 GAGTGATATTTTCAAGTCTTCTGCTGAGCT 1958

RESULT 5

BT009111 1632 bp mRNA linear PLN 20-JUN-2003
 LOCUS Triticum aestivum clone wll.pk0008.b3:fls, full insert mRNA
 DEFINITION
 sequence.
 ACCESSION BT009111
 VERSION BT009111.1 GI:32128662
 KEYWORDS FLI_CDNA.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 1632)
 AUTHORS Tingey, S.V., Walters, P., Powell, W., Dolan, M., Miao, G.-H.,
 Caraher, N.R., Hanafey, M.K. and Hainey, C.F.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-2003) Crop Genetics, E. I. Dupont de Nemours and
 Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
 USA
 FEATURES
 Location/Qualifiers
 1..1632
 source

The Salk, Stanford, POEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Banh, J., Bowser, L., Chan, M.M., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T., Karlin-Neuman, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers

```
1..1464
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U14970"
/ecotype="Columbia"
/notes="This clone is in pUNI 51"
```

CDS

```
1..1464
/notes="putative activator subunit of SNF1-related protein kinase SNF4"
/codon_start=1
/product="At1g09020/P7G19_11"
/protein_id="AA18191.1"
/db_xref="GI:23308443"
/translators="MFGSLDSSRMSAASQGLTPTFRVPMYGGRRVFLSGSTFRMT  
EHVMPGPEPTFOVINCITPGYHOYKEFDVGEWNRHDEHPFVSGGVWNTFIT  
QPWMPAGPESPTVGRNDVDVLEFADTADSOBAVRMSGVLELSRHRISVLLSTR  
TAYELLPSKSVLADVNLVYKQAFHLYEGLPLAPLWDFGQFVGLPGLDFILL  
LRSLGTHGSLTSEELHTTAAKWSQKHISQYDGRPYRPLVQVGFPLNKKDV  
ALXILQNKAAVPIVSSLDQSPQLHLASGLKICRYPRSSSLPILQOPI  
CSIPLGTVPIGESSKPLATLPRASLGSAALAIQAEVSIIPVDDNLSLIDYS  
RSDITLAKDKAQIHLDDTVHQALQGDASPPYGIENGQRCHMLRSDLSLVKM  
ERLANPQVRRLVIVEAGSKRVEGLISLSDVFQFLGL"
```

ORIGIN

```
Query Match 36.7%; Score 577.8; DB 8; Length 1464;
Best Local Similarity 66.4%; Pred. No. 1.1e-148;
Matches 901; Conservative 0; Mismatches 432; Indels 24; Gaps 4;

QY 5 GAGGGAACATTCGCGATCTCTATCGAAGGCTGCCACCTGCTATTTCAGGCTATTG 64
DB 129 GACAGAGCATGTGCCAATGTCACCGCTCGAGGGCTGCCCTACTCTTTTCAAGTCAT 188

QY 65 CAGTCTGTCTCTCGGGAATTCACAGTACAAATCTTTGTGACCGGGAAATGGCGCATGA 124
DB 189 CAATTGACGCCAGGATATCATCAGTATAGTTTGTGTGATGGGAATGGCGGCACGA 248

QY 125 TGAGCGTCAACCTACCATATCTCGGGAGTTTGGCATAGTTTAAACACITTTACTTGACAAG 184
DB 249 TGAGCACCAACCATTTGTAAAGCGGAATGGTGGAGTAGTGAATACAAATATTATTAAC 308

QY 185 GGAATATACCAATAAACACCTTATCAAGTCCAGCACACCTCGGAAGCAGATGAACAT 244
DB 309 GCCAGATA-----TGTTCTCTCTGCTTTTACGCCAGACACTGG---GCCGTCGAATAT 362

QY 245 GGATGTGGATAATGAAATTTTCAACGTCAGGTTTACGTTGTGTGATGGCAGCCGTTTCAGA 304
DB 363 GGATGTGATGATGTCTTCCTCGCA-----ACGGCTGACCCGTCGCCAGGA 407

QY 305 AGGTACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTCTGTTCTGAATA 364
DB 408 AGCTGTCTTAGGATGTGGGGGTGTGATTTGGAGCTTCTCGTCACCGCATATCGGTTT 467

QY 365 TCTGAATTTGCATACATGCTATGATTTTACTCCAGATTCTGGCAAGTATTATGGCCCTAGA 424
DB 468 ATTGTCAACCCGCACTGCATATGAGTGTCTCCCAAGATCGGCAAGTATTATGATTTGA 527

QY 425 CATTAATTACCTGTGAAGCAATCAITTCATATTCTCCATGAACAGGGGATTCCTGTAGC 484
```

```
DB 528 TGTAAATTCGCGAGTGAAGCAAGCATTCATATATCTCTATGAGCAGGAATCCCTTGGC 587
QY 485 TCCTCTCTCGGACTCATATTCAGAGGTCAATTTGTTGCTCTCTTAGCCCATTTGATTTTCAT 544
DB 588 TCCTCTTTGGGACTTTTGGAAAAGGCCAATTTGTTGAGTCTCTTGGTCCACTAGACTTCAT 647
QY 545 ACTCATATTTCGGGGNGCTAGAAACTCATGCTCGAATTCGACACTGACAGAGAGCTTGAAC 604
DB 648 TCTAAATCTCAGAGAGCTTGGAACTCATGATCAATCTTACAGAGGAAGCTTTGAGAC 707
QY 605 ACACACTATATCTGCATGGAAGAGGCTAAGCGGCCAACTAAATGGAAGAAATGATGTCA 664
DB 708 GCACACAATAGCAGCCCTGGAAAGGGAGGCTCATATTAGCCGCAATATGATGAAG 767
QY 665 GTGGCAGACCGAAGC---AGCATCTAGTGCATGACCCCTTATGATGCTCTTGGAGGACAT 721
DB 768 TGGGAGACCATATCTTAGGCCACTTTGTTCAAGTTGGACCCCTATGATTAATCTGAAGACGT 827
QY 722 TGCAGTAAAGCTTTTGCAAAATGCCATTTCTACAGTGCAGTATATTATTTATTCATCATCATC 781
DB 828 TGCCTGAAAATTTTGCAAAACAAGGTGGCAGCTGTTCCGGTTATATATTTCTTTTGCA 887
QY 782 AGATGATCATTTCCCGCAGTTATTGCACTTTTGCATCACTTTCTGGAATTTTGAATGAT 841
DB 888 GGATGGTTTCAATATCCGCAGTTACTGCACTTCTGCTGCTATCAGGCAATTTAAATATGAT 947
QY 842 TTGTAGATATCTTCAAAAATCAACTGTGTAATTTTGCCTATTCTTGAACCAACAGCTGTGCTC 901
DB 948 ATGCAGATACCTTTAGACATTCGCTAGCTCTTTTACCTATCTCTTTCAGAACCCATTTGTT 1007
QY 902 CATTCGCGCTGGTCTCTGGGTTCCGAAATTTGGTATCTGAAACAGTGTCTCCATTTGGCTAT 961
DB 1008 AATTCCTCTGGTACGTGGTCTCTAGAAATCGGGAATCAAGTAGCAAACTCTCTCGCTAC 1067
QY 962 GTTTCGACCTTAATTCCTCATCTTACCTCTGCTCTTAAACATGTTGTTTCAAGCTGGAGTAAG 1021
DB 1068 ATTGAGACCTTACGCTCTTTGGGTTCTGGGCTCGCATTTATTTTCAAGCTGAAGTCAG 1127
QY 1022 CTCAATACCAATTTGTGATGACAGGACTCCCTGCTTGGACATTTACTCTGAAGTGCAT 1081
DB 1128 TTCAATTCCTGCTAGTGGATGACAAACGACTCGCTATTGACATATCTCTCGAAGTATAT 1187
QY 1082 CACAGCCCTAGCAAAAGACAGGCTTACACACATGTTTCGGCTGGATGAGATGACCAATCA 1141
DB 1188 AACTGCTCTGGCTTAAAGCAAGGATACGACAGATTCATCTTGCATGACATGACAGTTCA 1247
QY 1142 CCAGGCTTTTGCAGCTTGGACAAGATGCCAATACACCTTTTGGATTTCTTTAAGGCCAGAG 1201
DB 1248 CCAGGCTTACAGTTGGGGCAAGATGCGAGCCGCTTATGGAATCTTCAACGGGCGAGAG 1307
QY 1202 ATGCCAGATGTGCTCGGCTCTGATCTTCTGCTGAAGTATGAGGCGACTGGGCTAATCC 1261
DB 1308 ATGTCACTAGTGTCTGCGCTCAGACTCTCTGTGAAAGTATGAGAGCGGTTGGCGAATCC 1367
QY 1262 TGGGCTGCGCGGGTGTTCATTTGTTGAAAGCTGGGAGCAAAAGCTGTGGAGGCAATCATATC 1321
DB 1368 AGGTGTAAGGAGGCTGGTGTATGTTGGAAGCAGGAGCAAAAGCTGTGAAAGTATCATTTTC 1427
QY 1322 ACTAAGTGACATTTTCAAGTTCTTGTGAGCTTTGTGA 1358
DB 1428 ACTGAGCGATGTTTTCCAATTTCTGCTCGGCTCTTTGA 1464
```

RESULT 7

AF250335

LOCUS

DEFINITION

protein kinase SNF4 mRNA, complete cds.

AF250335

VERSION

AF250335.1

KEYWORDS

Arabidopsis thaliana (thale cress)

2090 bp mRNA linear PLN 06-SEP-2000
Arabidopsis thaliana putative activator subunit of SNF1-related

ORGANISM	Arabidopsis thaliana	Db	1084	ATTGTCAACCCGCACTGTCATATGAGCTGCTCCAGAAATCGGCAAGCTATTATTCATTTGA 1143
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 2090)	Qy	425	CATTAACTTACCTGTGAAGCAATCATTCCTCATATTTCTCCATGAACAGGGATTCCTGTAGC 484
AUTHORS	Kleinow,T., Bhalarao,R., Brauer,F., Umeda,M., Salchert,K. and Koncz,C.	Db	1144	TGTTAATTTGCCAGTGAAGCAAGCATTCCTCATATCTATGACGAGGAATCCCTTTGGC 1203
TITLE	Functional identification of an Arabidopsis snf4 ortholog by screening for heterologous multicopy suppressors of snf4 deficiency in yeast	Qy	485	TCCTCTCTGGGACTCATTTGAGAGTCAATTTGTTGGTCTCTCTTAGCCCATTTGGATTTTCA 544
JOURNAL	Plant J. 23 (1), 115-122 (2000)	Db	1204	TCCTCTTTGGGACTTTGGAAAGAGCCAAATTTGTTGGAGTTCTTTGGTCCCTAGACTTCAT 1263
MEDLINE	20367008	Qy	545	ACTCATATTTGCGGAGCTAGAACTCATGCTCGAACTTGCACAGAGAGAGCTTGAAC 604
PUBMED	10929106	Db	1264	TCTAATACTGAGAGAGCTTGGAACTCATGATCAAACTTGCACAGAGAGAGCTTGAAC 1323
REFERENCE	2 (bases 1 to 2090)	Qy	605	ACACATATCTGCAATGGAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATCTCA 664
AUTHORS	Kleinow,T., Bhalarao,R. and Koncz,C.	Db	1324	GCACAATAGCAGCTTGGAAAGAGGGAGGCTCATATTAGCCGCAATATGATGGAAG 1383
TITLE	Direct Submission	Qy	665	GTGGGACCGCAAC---AGCATCTAGTGCATGCCACCCCTTATGAGTCCTTTGAGGGACAT 721
JOURNAL	Submitted (29-MAR-2000) Genetic Principles of Plant Breeding, Max-Planck-Institute for Plant Breeding, Carl-von-Linne-Weg 10, Koeln 50829, Germany	Db	1384	TGGGAGACCATATCTCTAGGCCACTTGTTCAGGTTGGACCCCTATGATAATCTGAAAGAGCT 1443
FEATURES	Location/Qualifiers	Qy	722	TGCAGTAAGCTTTTGCAGAAATGGCAATTTCTACAGTGCAGATTTATTTATTCATCATCTC 781
source	1..2090	Db	1444	TGCCCTGAAAAATTTTGCAGAAACAAAGGTGGCAGCTGTTCCGTTTATATATTAATTCCTTTGCA 1503
CDS	/note="contains 4 cystathione-beta-synthase (CBS) domains; heterologous multicopy suppressor of snf4 deficiency in yeast"	Qy	782	AGATGATCATTTCCCGCAGTTATTGCATCTTGCATCACTTCTGGAATTTTGAATATGAT 841
	/codon_start=1	Db	1504	GGATGTTTCAATATCCGAGTTACTGCACTCTTGCTTCTGCTATCAGGCATATTAATAAUGTAT 1563
	/product="putative activator subunit of SNF1-related protein kinase SNF4"	Qy	842	TTGTAGATATCTTCAAAACTCAACTGCTAATTTTGCCTTATTTCTGAACCAACAGTGTGCTC 901
	/protein_id="AAG10141.1"	Db	1564	ATGCAGATATCTTTAGCAATTTGCTAGCTCTTTTACCTATCTCTTCAGCAACCCATTTGCTC 1623
	/db_xref="GI:9365729"	Qy	902	CATTCCGCTGGGTTCTCGGTTCCGAAATATGTTGATCTGAAACAGTCTCATTTGGCTAT 961
	/translation="MVPAGFSPETIGRSNMVDVFLRTADPSQERAVPMRGVDLELS RHRIYVLLSTRYELLPESSGYIALVDNLVVKQAFHLYEQIPLAPLWDFGKGQFV GVLPDLFTLIURELTHGSNLTETEEHTTAAWKEGKAHISROYDGSGRPYRPLV OYGPYDNLKVALKILQNKVAAVPVITYSLODGSYPOLLHLASLSGLKICIRPHS SSPLPILQPCISPLGTWVPRIGSSSKPLATLRPHASLSGNALLVQAEVSSIPVY DNDLSLDIYSIDITALAKAYAQIHDDMFVHQALSLGQDASPPYGFNGQRCMH CLRSLSLVKVMERLANPVRRLVIVEAGSKRVEGHIISLSDVFQFLGL"	Db	1624	AATTTCCCTGGGTACGTGGGTTCTAGATCGGAGAAATCAAGTAGCAAACTCTTTGCTAC 1693
		Qy	962	GTTCGACCTAATGSCCTCACTAGCTAGCTGCTCCCTTAACTGTTTGGTTCAAGCTGAGTAAG 1021
ORIGIN		Db	1684	ATTGAGACCTCAGCCCTCTTTGGGTTCTGGGCTCGCATTAATTAATTAATTAATTAATTAAT 1743
Query Match	36.7%; Score 577.8; DB 8; Length 2090;	Qy	1022	CTCAATACCAATTTGTGATGACAAACGACTCCCTCTGTTTGACACTTACTCTAGAAGTGCAT 1081
Best Local Similarity	66.4%; Pred. No. 1.1e-148;	Db	1744	TTCAATTTCCGTTAGTGGATGACAAACGACTCGCTTATTTGACATATACCTCTCGAAGTGATAT 1803
Matches	901; Conservative 0; Mismatches 432; Indels 24; Gaps 4;	Qy	1082	CACAGCCCTAGCAAAAGACAGGTCTACACAGATGTTGGCTGATGAGATGACCATCA 1141
Qy	5	Db	1804	AACCTGCTTGGCTTAAAGACAAAGGCATACGACAGATTCATCTTGATGACATGACAGTTCA 1863
Db	745	Qy	1142	CCAGGCTTTGACGCTTGGCAAGATGCCAATACACCTTTTGGATTTCTTTTAAACGCCACAGAG 1201
Qy	65	Db	1864	CMAGGCTTACACTTGGGCAAGATGCGAGCCGCTTATGGAATCTTCAACGGGACAGAG 1923
Db	805	Qy	1202	ATCCAGATGTGCTCCGTTCTGATCCTTTGCTGAAGGTGATGAGAGCACTGGCTAATCC 1261
Qy	125	Db	1924	ATGTCACATGTGCTTGGCTCAGACTCTCTTTGTTGAAAGTGTATGAGGCGGTTGGCGAATCC 1983
Db	865	Qy	1262	TGGGTTGGCGCGGTTGTTGATTTGGAAGCTGGGAGCAACGTTGTGAGGCGCATCATTC 1321
Qy	185	Db	1984	AGGTGTAAGAGGCTGGTGATAGTGAAGCAGGAGCAACGTTGTGAAGGTATCATTTTC 2043
Db	925	Qy	1322	ACTAAGTGACATTTTCAAGTTCTTGTGAGCTTGTGA 1358
Qy	245	Db	2044	ACTGAGGATGTTTCCAAATTCCTGCTCGTCTTTGA 2080
Db	979	Qy		
Qy	305	Db		
Db	1024	Qy		
Qy	365	Db		

RESULT 8
 AF439826
 LOCUS
 DEFINITION Arabidopsids thaliana At:g09020/F7G1.9_11 mRNA, complete cds.
 ACCESSION AF439826

AF439826.1 GI:16612254
 FLZ CDNA
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 2301)
 Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Koesema, E., Meyers, M.C.,
 Banh, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D.,
 Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A.,
 Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X.,
 Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
 Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A.,
 Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
 Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Arabidopsis cdna clones
 Unpublished
 2 (bases 1 to 2301)
 Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Koesema, E., Meyers, M.C.,
 Banh, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D.,
 Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A.,
 Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X.,
 Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
 Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A.,
 Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
 Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Direct Submission
 Submitted (24-OCT-2001) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFI cDNAs (RAFI CDNA : 'RIKEN
 Arabidopsis Full-length cDNA' : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, FGENC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFI cDNAs: Shinn, P., Chen, H.,
 Cheuk, R., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bowser, L.,
 Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,
 Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S.,
 Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M.,
 Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W.,
 Theologis, A., and Ecker, J.R.
 Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.
 Location/Qualifiers
 .. 2301
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="RAFL09-86-B11 (R14970)"
 /ecotype="Columbia"
 1..603
 504..2067
 /note="putative activator subunit of SNF1-related protein
 kinase SNF4"
 /codon_start=1
 /product="At1g09020/F7G19.11"
 /protein_id="AAL27498.1"
 /db_xref="GI:16612255"
 /translation="MFGSTLDSRGNSAAGSGLLTPTRFVWPGYGRVFLSGFTRWIT
 EHPVMSLEGCTVQVINCILNLTGPHQXKFVDGWRHDEHQPFGVSGNGVYVITPIT
 GPDVMPAGSPETLGRSNMVDVFLRTADPSQEAAPVPMGSDVLELSRRHSVLLSTR
 TAYELLPSGKIVLADNLVPRQAFHLYEQLPLAPWDFGKGFVGLVGPDLFILI
 LRELTHGNSKTEELEYHTIAAKGKRAHISROYDGSGRPYPRFLVQVPGDNLKDV
 ALKILQNVAAVPVLYSSIDGSPYQLHLASLGLKICRYFRHSSSLPILOQPI
 CSIPLGTWPRIGSSSKPLATLRPHASLGSALALLVQAEVSSIIPWDNDLSLIDIS

misc_difference 1267
 /note="compared to genome"
 /replace="g"
 2068..2301

3'UTR
 ORIGIN

Query Match 36.7%; Score 577.8; DB 8; Length 2301;
 Best Local Similarity 66.4%; Pred. No. 1.1e-148;
 Matches 901; Conservative 0; Mismatches 432; Indels 24; Gaps 4;

Qy 5 GAGGGAACATTTGCGGATGCTCTCTATCGAGGCTGCCCACTGATTTTCAAGGCTATTG 64
 Db 732 GACAGAGATGTGCAATGTCCCGCTCGAGGCTGCCCTCTCTCTCTCTCTCTCTCTCT 791
 Qy 65 CAGTCTGCT 124
 Db 792 CAACCTTGAGCCAGGATATCATCATAGTATAGTATTTTGTGATGGGAATGGCGCAGA 851
 Qy 125 TGAGGCTCAACCTACCATATCTGGGAGTTTGGCATAGTTTAAACACATTTACTTTGACAAG 184
 Db 852 TGAGCACCACCACTTTGTAGCGGGAATGGTGGAGTAGTGAATACAAATTTTATAACTGG 911
 Qy 185 GGAATATAACCAATAAACAACCTTATCAAGTCCCAAGCACACTGGAAGCAGGATGAACAT 244
 Db 912 GCCAGATA---TGCTTCT 965
 Qy 245 GGATGTGGAATGAATAATTTTCAACGCTACGGTTACGTTGTGAGATGACACCGTTTCAGA 304
 Db 966 GGATGTGGAATGATGTTCT 1010
 Qy 305 AGGTACTCTGAGAGTTTCAGAGGCTGCATACAAATATCTAGGTGCTGTTTCTGAATA 364
 Db 1011 AGCTGTTCTCTAGATGTCGGGGTGTGATTTGGAGCTTTCTCTCTCACCGATATCGGTTT 1070
 Qy 365 TCTGAATTTGCATACATCATGATTTTACTCCAGATTCGCGCAAGTTTATTCGCTTACG 424
 Db 1071 ATTGTCAACCGCACTGCATATGAGTCTCCAGAAATCGGCAAGTTTATTCGATTGGA 1130
 Qy 425 CATTAATTTACCTGTGAAGCAATCATTCATTTCCATGAAAGAGGATTCCTGTAGC 484
 Db 1131 TGTAAATTTGCCAGTGAAGCAAGCAATTCATATCTCTATGACGAGGAATCCCTTTGGC 1190
 Qy 485 TCCTCTCTGGAACCTCAATTCAGAGTCAATTTGTTGCTCTCTTACGCAAGTTGATTTTCA 544
 Db 1191 TCCTCTTTGGGACTTTGGAAAAGGCCAATTTGTTGGAGTTCTTGTCTCCTATGAGCTTCT 1250
 Qy 545 ACTCATATTGGGAGCTAGAAACTCATGGCTCGAACTTGACAGAGAGCAGCTTGAAC 604
 Db 1251 TCTAATCTGAGAGAGCTTGGAACTCATGTGATCAAACTTGACAGAGAGAGCTTGAGAC 1310
 Qy 605 ACACACTATATCTGATGGAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATAGTCA 664
 Db 1311 GCACACAATAGCAGCTCGGAAAGAGGGAAGGCTCATATTAGCCGGCAATATGATGGAG 1370
 Qy 665 GTGGGCGCGCAAC---AGCATCTAGTGCATGCCACCCCTATGAGTCTCTGAGGAGCAT 721
 Db 1371 TGGGAGACCATATCTCTAGGCCACTTGTTCAGTTGGAACCTATGATATCTGAAGAGCT 1430
 Qy 722 TGCACTAAAGCTTTTSCAAAATGGCAATTTCTACAGTGGCAGTATTTATTCATCATCATC 781
 Db 1431 TGCCCTGAAAATTTTGGAAAACAAGGTGGCAGCTGTTCCGGTTATATATTTCTTCTTGA 1490
 Qy 782 AGATGATCATCTCCGAGTTATTCATCTCTGATCATCTCTCTGAAATTTTGAATATGAT 841
 Db 1491 GGATGGTTTATATCCGAGTTTACTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1550
 Qy 842 TTGTAGATCTTCAAAAACCTCAACTGGTAAATTTGCTTATTTCTGAACCAACCAAGTGTG 901
 Db 1551 ATGCAGATCTTTAGACATTCCTCTAGCTCTTTTACCTATCTCTCTCTCTCTCTCTCTCT 1610
 Qy 902 CATTCGCTGGGTTCTCTGGGTTCCGAAATTCGGTATCTGAACAGTCTGCTTCTGCTAT 961

5'UTR
 CDS

FEATURES
 source

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

COMMENT

Db	1073	CCTTTGGGTACATGGGTGGCTAAACGTTGGGATCCAAATGGACAGCCACTGATAAGGCTA	1132	Db	132	CGTTTCGGCAGCTGTTGCACTCTTGGCATCCCTTTCCAGGAATTTTGAATGTATCTGTAGAT	191
Qy	966	CGACCTAATGCCCTCACTTAGCTCTGCGCTTAAACATCTGTTGGTTCCTCAAGCTGGAGTAAGCTCA	1025	Qy	850	ACTTCAAAACTACACTGTTGTAATTTGGCTATTCTGAACCAACAGTGTGCTCCATTCGGC	909
Db	1133	AGGCCAAATGCTTCTCTTGGTGATGCCCTGTCGATGTTGTTCAAGCTAAAGTCAGCTCA	1192	Db	192	ACTTCAAGAACTCCACTGTTGCGGATTTTGGCGAATTTAAACCAACAGTAGTGTCAATTCGGC	251
Qy	1026	ATACCAATTTGGATGACCAAGCACTCCCTGCTTGACACTTACTCTAGAAATGATGATCACA	1085	Qy	910	TGGGTTCTCGGGTTCGGAATTTGGTGATCTGAACAGCTGCTCCATTCGGCTATGTTGGCAC	969
Db	1193	ATACCAATTTGGATGAAATGATTCATCTGCTTGACATATATTCAAGAGTGATTTACT	1252	Db	252	TGGGTACTCGGGTTCGGAATTTGGTGACCAAAATTTGGTGACCAAAATGCTATTCATTCGGCTATGTTGGGC	311
Qy	1086	GCCCTAGCAAAAGACAGGCTCTACACATGTTCCGGCTGGATGAGTACCATTCACACAG	1145	Qy	970	CTAATGCTCCTCCTTAGCTCTGCGCTTAACTGTTGGTTCAGCTGGAGTGAAGCTCAATAC	1029
Db	1253	GCATTTGGCTAAGATAAGCATATGCTCGATATCCCTGGATGAACTAATATTCATCAG	1312	Db	312	CTAATACATCTCTTAGCTCTGCGCTTAACTGTTGGTTCAGCTGGAGTGAAGTCAATAC	371
Qy	1146	GCTTTGCACTTGGACAAAGATGCCAATACACTTTTGGATCTTTTAAAGCCGACAGATGC	1205	Qy	1030	CAATTTGGATGACAGCACTCCCTGCTTGACACTTACTCTAGAAATGATGATCAGACCC	1089
Db	1313	GCATTAATTTGGGCAAGATGCAATCTCTTATGCTTTAACAATGGGACAGATGT	1372	Db	372	CCATTTGGATGATTAAGCACTCGCTGATCGACACATATCCAGAAATGATCAGACGCTC	431
Qy	1206	CAGATGTCCTCGGCTGATCTCTTTGCTGAAGTGTAGAGCACTGGCTAATCTCGG	1265	Qy	1090	TAGCAAAAGACAAAGTCTTACACATGTTCCGCTGGATGAGATGACCAATTCACAGGCTT	1149
Db	1373	CATATGTTGAGGCTGATTCATTCACAAAGTGTAGAGCGGTTGGCTAAACCTGGG	1432	Db	432	TAGCAAAAGACAAAGTCTTACACATGTTCCGCTGGATGAGATGACCAATTCACAGGCTT	491
Qy	1266	GTGCGGGGGTGTTCATGTTGGAAGCTGGAGCAACGTTGGAGGGCATCATATCACTA	1325	Qy	1150	TGCAGCTTGGACAAAGATGCCAATACACTTTTGGATCTTTTAAAGCCGACAGATGCCAGA	1209
Db	1433	GTTAGGAGACTGTGATTTGGTGAAGCTGGTACGAGCGTGTGGAGGGATTTTTCATTA	1492	Db	492	TGCAGCTTGGACAAAGATGCCAATACACTTTTGGATCTTTTAAAGCCGACAGATGCCAGA	551
Qy	1326	AGTGACATTTTCAAGTCTTCTGCTGAGCTTTTGAGGAGAGA	1365	Qy	1210	TGTCCTCGGCTCTGATCTTCTGCTGAGGCTGGAGGCTGCTTAATCTCGGCTGC	1269
Db	1493	AGTGATCTTTTCAAGTCTTCTGCTGAGCTTTTGAGGAGAGA	1532	Db	552	TGTCCTCGGCTCTGATCTTCTGCTGAGGCTGGAGGCTGCTTAATCTCGGCTGC	611
RESULT 10				Qy	1270	GGCGGGTTCATTTGTTGGAAGCTGGAGCAACCTGTGGAGGGCATCATCATCACTAAGTG	1329
BT009623				Db	612	GTCGCTGTTTCATCTGCTGAGGCTGGAGGCTGGAGGCTGCTTAATCTCGGCTGC	671
LOCUS				Qy	1330	ACATTTTCAAGTCTTCTGCTGAGCTTTGCA	1358
DEFINITION				Db	672	ACATATTCAAGTCTGCTGAGCTAGCGA	700
ACCESSION				RESULT 11			
BT009623				AK1211122			
VERSION				LOCUS			
BT009623.1				DEFINITION			
FLI CDNA.							
SOURCE				ACCESSION			
Triticum aestivum (bread wheat)				AK1211122			
ORGANISM				VERSION			
Triticum aestivum				AK1211122.1			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				KEYWORDS			
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				FLI CDNA; CAP trapper.			
Poideae; Triticeae; Triticum.				SOURCE			
1 (bases 1 to 982)				Oryza sativa (japonica cultivar-group)			
Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,				ORGANISM			
Caraher,N.R., Hanafey,M.K. and Hainey,C.F.				Oryza sativa (japonica cultivar-group)			
Direct Submission				Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and				Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,				Ehrhartioideae; Oryzeae; Oryza.			
USA				REFERENCE			
Location/Qualifiers				1			
1..982				AUTHORS			
/organism="Triticum aestivum"				The Rice Full-length cDNA Consortium, National Institute of			
/mol_type="mRNA"				Agrobiological Sciences Rice Full-length cDNA Project Team:-			
/db_xref="taxon:4565"				Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,			
/clone="wreln.pk0143.e2.fis"				Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,			
				Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,			
				Ohtsuki,K., Shishiki,T., Foundation of Advancement of International			
				Science Genome Sequencing & Analysis Group: Ohtsuki,Y., Murakami,K.,			
				Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,			
				Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,			
				Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,			
				Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,			
				Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,			
				Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,			
				Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,			
				Kagawa,I., Kondo,S., Konno,K., Miyazaki,A., Osato,N., Ota,Y.,			
				Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,			
				Yoshino,M. and Hayashizaki,Y.			
				Collection, mapping, and annotation of over 28,000 cDNA clones from			
				Japonica rice			
				JOURNAL			
				Science 301 (5631), 376-379 (2003)			

22752273	ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodana, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Ito, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Nemiki, T., Narioka, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.
12869764	Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice
2	Unpublished 3 (bases 1 to 1923)
	Kikuchi, S.
	Direct Submission
	Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: sikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)
	This clone is one of the 32K full-length cDNA clones from japonica rice.
	URL: http://cdna01.dna.affrc.go.jp/cDNA/
	NTAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M. and Nakahama, Y.
	FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodana, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narioka, R., Niikura, J., Oka, M., Ryo, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
	Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Haragaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kouda, M., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ooka, Y., Saitoh, H., Sakai, C., Sakai, K., Sasaki, D., Sato, K., Shibata, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sogabe, Y., Tagami, M., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
	Location/Qualifiers
	1. 1923
	/organism="Oryza sativa (japonica cultivar-group)"
	/mol_type="mRNA"
	/cultivar="Nipponbare"
	/db_xref="taxon:39947"
	/clone="J023074117"
FEATURES	source
ORIGIN	

Db 1825 1826

RESULT 12
AC128647
LOCUS
DEFINITION
AC128647 92392 bp DNA linear PLN 04-SEP-2003
Oryza sativa chromosome 3 BAC clone OSUNB0062G19, complete
sequence.
VERSION
AC128647.8 GI:34447243
KEYWORDS
HTG.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartidae; Oryzeae; Oryza.
1 (bases 1 to 92392)
Buell, R., Hsiao, J., Zismann, V., Moffat, K.M., Hill, J.,
Gansberger, K., Burgess, S., Jarrahi, B., Shvartsbeyn, M., Brenner, M.,
Ciecko, A., Pai, S., Vanaken, S., Hansen, C., Utterbach, T.,
Feldblyum, T., Khalak, H.G., Yuan, Q., Quackenbush, J., White, O.,
Salzberg, S. and Fraser, C.
Oryza sativa ssp. japonica cv. Nipponbare OSUNB0062G19 BAC genomic
sequence
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 92392)
Buell, R.
Direct Submission
Submitted (20-JUL-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
JOURNAL
3 (bases 1 to 92392)
Buell, R.
Direct Submission
Submitted (10-APR-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
JOURNAL
4 (bases 1 to 92392)
Buell, R.
Direct Submission
Submitted (04-SEP-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
JOURNAL
COMMENT
On Sep 4, 2003 this sequence version replaced gi:28827864.
FEATURES
source
1. 92392
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSUNB0062G19"

ORIGIN
Query Match 8.6%; Score 135; DB 8; Length 92392;
Best Local Similarity 64.4%; Pred. No. 5.5e-26;
Matches 266; Conservative 0; Mismatches 65; Indels 82; Gaps 1;
79838 TATTGTTAGGCCACCCCTTATGAATCCCTGAGGGAATTCGTATGAAGTACTGCAAAA 79897
79839 TCGATGTCATGCGCCCTTATGAGTCTCTGAGGACATTCGAGTAAGCTTTTGCAGAA 742
79840 TGGCATTTCTACATGCGCAGTTATTTATCATCATCATCATCATCATTCCTCCGCGATT 802
79898 CGGTGTTCTACCGTTCCAAATATGTTTCTCATCCAGATGCTCATATCCACAATT 79957
803 ATTGCATCTTGCATCAGCTTCTTGGAAATTTGAAAT----- 837
79958 GTTGCATCTTCTCCCTTCTGGAATTTTGAATGTAAGTGTTCATGATTTCATG 80017
838 -----GTA 840
80018 AAAAGTGTGTATACAAACATATTATCATGATAGCTTAAACAATTTCAATTTTCAGGTA 80077

QY 841 TTTGTAGTACTTCAAAACTCACTGTAATTTGCTATTTCTGAACCAACCGTGTCT 900
Db 80078 TTTGTAGTATATTTTAAAAATTTCTCAAGTAATTTTACCTATTTTGAAGCAACTGTATGCA 80137
QY 901 CCATTCCGCTGGGTTCCTGGTTCGGAAATTTGGTATCTGAACAGTCTGTCATTTGGCTA 960
Db 80138 CAATTCCTCTGGTACCTGGTTCACAAATTTGTATCTCTATATGCGCGTCCATTGGCTA 80197
QY 961 TGTTCGACCTTAATGCTCTACTTACCTCTGCTTACCAATTTGTTGTTCAAGCT 1013
Db 80198 TGTTCGCGCTTACACATCTCTTAGCGTTCGCTGAAATTTGCTGTTCAAGT 80250

RESULT 13
AC096690/c
LOCUS
DEFINITION
AC096690 145115 bp DNA linear PLN 25-APR-2003
Oryza sativa chromosome 3 BAC OSUNB0059G06 genomic sequence,
complete sequence.
AC096690
VERSION
AC096690.5 GI:29568036
KEYWORDS
HTG.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartidae; Oryzeae; Oryza.
1 (bases 1 to 145115)
Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Gansberger, K., Jones, K.M.,
Overton II, L.L., Tsitrin, T., Kim, M.H., Bera, J.J., Jin, S.S.,
Fadrosh, D.W., Tallon, L.J., Koo, H., Zismann, V., Hsiao, J., Blunt, S.,
Vanaken, S.S., Riedmuller, S.B., Utterbach, T.T., Feldblyum, T.V.,
Yang, Q.Q., Haas, B.J., Suh, B.B., Peterson, J.J., Quackenbush, J.,
White, O., Salzberg, S.L. and Fraser, C.M.
Oryza sativa chromosome 3 BAC OSUNB0059G06 genomic sequence
Unpublished
REFERENCE
2 (bases 1 to 145115)
Buell, R.
Direct Submission
Submitted (22-SEP-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
JOURNAL
3 (bases 1 to 145115)
Buell, R.
Direct Submission
Submitted (08-MAY-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
JOURNAL
4 (bases 1 to 145115)
Buell, R.
Direct Submission
Submitted (05-APR-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
On Apr 5, 2003 this sequence version replaced gi:20503077.
Address all correspondence to:rice@tigr.org

BAC clone OSUNB0059G06 is from Oryza sativa chromosome 3
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (<http://www.softberry.com/>),
GENSCAN and GENSCAN (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHMM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/GeneMark/>), and Geneslicer
(Mihaila Perlea and Steven Salzberg, contact mperlea@tigr.org),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (<http://www.tigr.org/tdb/cgi.shtml>).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as unknown proteins.

Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

This BAC overlaps with rice BAC OSUNBB0062G19 (AC128647) and OSUNBa0033P04 (AC092263).

FEATURES

source

```

1. .145115
  /location="Qualifiers
  /organism="Oryza sativa (japonica cultivar-group)"
  /mol_type="genomic DNA"
  /cultivar="Nipponbare"
  /db_xref="taxon:39947"
  /chromosome="3"
  /map="near 1375"
  /clone="OSUNBa0059G06"
  /note="japonica cultivar-group"
  1271..1331
    /rpt_family="AT-rich"
    2342..3105
      /gene="OSUNBa0059G06.1"
      /note="predicted by genemarkEMM"
      join(2342..2619,2999..>3105)
      /gene="OSUNBa0059G06.1"
      join(<2342..2619,2999..3105)
      /gene="OSUNBa0059G06.1"
      /codon_start=2
      /product="hypothetical protein, 5'-partial"
      /protein_id="AAP21388.1"
      /db_xref="GI:30102982"
      /translation="EQWRAEAGGGHRCRGWRARERCOLREGVEAAGGGRARRG
      RGLCHRRHTAASSSHVARELPPELSLNLEFLSSWWSASAVDPQASGAUNSE
      RLHRAASAAATARSASATVGR"
      2353..2479
        /rpt_family="G-rich"
        3192..3467
          /rpt_family="(CCG)n"
          5230..5343
            /rpt_family="(TATAA)n"
            5868..8753
              /gene="OSUNBa0059G06.2"
              /note="similar to family II lipase EXL3 GB:AAK30018
              GI:15054386 (Arabidopsis thaliana); contains Pfam profile
              PF00657 (GDSL like lipase Acylhydrolase)"
              join(<5868..6150,7939..>8753)
              /gene="OSUNBa0059G06.2"
              join(5868..6150,7939..8753)
              /gene="OSUNBa0059G06.2"
              /codon_start=1
              /product="putative lipase/acylhydrolase"
              /protein_id="AAP21383.1"
              /db_xref="GI:30102970"
              /translation="WTMKKQKKNNKGLLVAVAVAMAPAGSPASSGRCAPMPAV
              IVGDSIVDPNNNLTQIKANHAPYGMDFANSEPTGRYSNGLIPTDFIVQGLNVKQ
              LMPYLGVLSPELKTGVSPASCATGVPLTPVIVSVITLDDQOIEFYHVRKRLGV
              VGEETARIDGALVVCAGCTDDIANTYFTTPFRSEVYDIPSYVDLLVSGNAKLDDV
              AAGLRIGIFGPPICVPSQRTLGQGGPHRRCKERNYAAKLFNSRMEVEIAKTNP
              ATTMVVDYITLQELVENGDKYFTETTRCGGTGTLVETGLCDARFVDCIDNVSN
              HVFSDSYHPTQRAYKIIVDYIFDNYLQFLA"
              6410..6437
                /rpt_family="AT-rich"
                7781..7804
                  /rpt_family="(CGAT)n"
                  12122..113806
                    /gene="OSUNBa0059G06.3"
                    /note="predicted by fgenesh"
                    join(<12122..12162,12851..13020,13702..>13806)
                    /gene="OSUNBa0059G06.3"
                    join(12122..12162,12851..13020,13702..13806)
                    /gene="OSUNBa0059G06.3"

```

```

/codon_start=1
/product="hypothetical protein"
/protein_id="AAP21385.1"
/db_xref="GI:30102972"
/translation="MRVRKGVGASPGFKESVSGAGWNGCAPRAGDERPPPGADGNG
SEMPIDGGARKWFGVGLGSSRTVPARVDELGPDDARNMAMAMWAAKAVDCTAEF
SSREGLARGSGESWQPCSELAAMRRQSMITVAEVELDSGNRESRDKFKVDSDIRLIG
TSIMRRFKGGD"
13321..13364
  /rpt_family="GA-rich"
  15008..15049
    /rpt_family="(TC)n"
    16325..16352
      /rpt_family="(TTCC)n"
      16964..17005
        /note="Overlap BAC OSUNBa0033P04 has 49 TA repeats in this
        region."
        16964..17005
          /rpt_family="(TA)n"
          17177..17203
            /rpt_family="AT-rich"
            17384..18823
              /gene="OSUNBa0059G06.4"
              /note="contains Pfam profile PF01554 (MatE)"
              <17384..>18823
                /gene="OSUNBa0059G06.4"
                17384..18823
                  /gene="OSUNBa0059G06.4"
                  /codon_start=1
                  /product="putative MatE efflux family protein"
                  /protein_id="AAP21388.1"
                  /db_xref="GI:30102975"
                  /translation="MTDPPSPPHERTKWAESVASEFRAQGIAPFLAKMLTWFAL
                  AVTTAFLGRGLQLAAGTLGFSFANVTGFAVLTLGLCAAMDPICGQAHGASNGKLARK
                  TLWATFILLGASTPIAFPLHVDVALLRSCQADMSNABSVVCLLPDLAVTSEVK
                  ELKSLSAQGVITPLFASALALALHVLPLNWKARTGIIQSVATAVWVSDLAVALA
                  GYLVSRERRKAGGGGVQETRGWVLLRLAVPSCINTCLEWVCYEILVLLGRLP
                  DARRTAVMAVTLLNFDYLLFAGMLSLSVASVRSVNSBELGAGEAARRAGWVSIVGGA
                  VGGVGGGVAAVARRAMGSISSDAGVREGVGRAMVAVLEVNFPLNVCQGIVRGT
                  ARPAGVMYAVAGFYVALPLGVALAFKARLIGIQGLLLGLFVGAAASLAVLLTFIARM
                  DWPAAQAKARTTATVAQFHHQDEVVQP"
                  18325..18515
                    /rpt_family="G-rich"
                    20002..21960
                      /gene="OSUNBa0059G06.5"
                      /note="EST CA759271, BE040176 from this gene"
                      join(<20002..20123,20242..20298,20375..20500,20579..20654,
                      20811..20903,21132..21245,21366..21960)
                      /gene="OSUNBa0059G06.5"
                      join(20002..20123,20242..20298,20375..20500,20579..20654,
                      20811..20903,21132..21245,21366..21419)
                      /gene="OSUNBa0059G06.5"
                      /codon_start=1
                      /product="unknown protein"
                      /protein_id="AAP21391.1"
                      /db_xref="GI:30102978"
                      /translation="MAPPPCCVLVVRCCSLSPSPPPPPSPNNLRARREFF
                      VNSPDLALAIPLHSGTKPRLQVQKSPSPSLALCPATNNCVSTCEDITDSIHVAPPW
                      LRSKDKRRAKPIKHEAINQIOWTCTKEDNFTPLVKTDDYRVEYESPIPGFV
                      DDVEFPFGNKSIVQYRSASRSRGSFIDENANKVKVRLALEKNWASESTI"
                      20063..20084
                        /rpt_family="(CCG)n"
                        20673..20703
                          /rpt_family="(TTTTTC)n"
                          22163..22193
                            /rpt_family="AT-rich"
                            24255..25676
                              /gene="OSUNBa0059G06.6"
                              /note="predicted by fgenesh"

```

Query Match 8.6%; Score 135; DB 8; Length 145115;
 Best Local Similarity 64.4%; Pred No. 5.5e-26;
 Matches 266; Conservative 0; Mismatches 65; Indels 82; Gaps 1;

```

Qy 683 TCTAGTCATGCCACCCCTTATAGTCTTCTGAGGGACATTCAGCAATTCGTTTGCAGAAA 742
Db 135226 TATTTGTTAGGCCACCCCTTATGAATCCCTGAGGGAAATTCCTATGAAGATCTGCAAAA 135167
Qy 743 TGGCATTTCTACATGSCAGTTATTTATTCATCATCATCAGATGATCATTCGCCGAGTT 802
Db 135166 CGGTGTTTCTACCGTTCOAATATGTTTCTCATCACCAGATGGCTCATATCACAAT 135107
Qy 803 ATTGCATCTTGCATCAGCTTCTCGAAATTTGAAAT----- 837
Db 135106 GTTGCATCTTGCCTCCCTTCTCGAAATTTGAAATGTAAGTGTCTATGATTTGCAATG 135047
Qy 838 -----GTA 840
Db 135046 AAAAGTGTCTATACAAACATATTCATGATAGCTTTAACAACATTTTCATTTTCAGGTA 134987
Qy 841 TTCTGATATCTTCAAAAACCTCAACTGTATTTTGCCTATTCTGAACCAACAGCTGTCT 900
Db 134986 TTTGTAGATATTTAAAAATTTCTCAAGTAAATTTACCTATTTTGGCCACACTGTATGCA 134927
Qy 901 CCATTCCGCTGGTCTCCGTGCTCGGAAATTTGGTGAATCGAACAGTCTCCATTTGGCTA 960
Db 134926 CAATTCCTCTGGTACCTGGTTCGAAATTTGGTGAATCTCAATGCGCTCCATTTGGCTA 134867
Qy 961 TGTTCGACCTAATGCTCCTACCTAGCTCTGCGCTTAACATGTTGGTTCAAGCT 1013
Db 134866 TGTTCGCGCTTAACACATCTCTTAGCGCTCGCTCGAATTTGCTGTTCAAGGT 134814

RESULT 14
AC146552/c
LOCUS AC146552 107857 bp DNA linear HTG 30-AUG-2003
DEFINITION Medicago truncatula clone mth2-9d1, WORKING DRAFT SEQUENCE, 9
unordered pieces.
ACCESSION AC146552
VERSION AC146552.1 GI:34365866
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 107857)
Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
Medicago truncatula BAC Clone mth2-9d1
Unpublished
2 (bases 1 to 107857)
Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
Direct Submission
Submitted (30-AUG-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT ----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code: UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 3029: contig of 3029 bp in length
* 3030 3129: gap of unknown length
* 3130 7573: contig of 4444 bp in length
* 7574 7673: gap of unknown length

```

```

* 7674 16205: contig of 8532 bp in length
* 16206 16305: gap of unknown length
* 16306 26640: contig of 10335 bp in length
* 26641 26740: gap of unknown length
* 26741 36309: contig of 9569 bp in length
* 36310 36409: gap of unknown length
* 36410 48365: contig of 11956 bp in length
* 48366 48465: gap of unknown length
* 48466 60592: contig of 12126 bp in length
* 60592 60692: gap of unknown length
* 60692 76823: contig of 16132 bp in length
* 76824 107857: contig of 30934 bp in length.
* 76924 107857: contig of 30934 bp in length.
FEATURES             Location/Qualifiers
     source            1..107857
     organism="Medicago truncatula"
     mol_type="genomic DNA"
     db_xref="taxon:3880"
     clone="mth2-9d1"
     clone_lib="Medicago truncatula BAC library H2"

ORIGIN
Query Match          5.5%; Score 87; DB 2; Length 107857;
Best Local Similarity 63.8%; Pred. No. 1.1e-12;
Matches 132; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 681 CATCTAGTCATGCCACCCCTTATGAGTCTTCTGAGGGACATTCAGCAATTCGTTTGCAG 740
Db 94133 CTTATTGTGCGAGCGGCCCATCTGATAATTTGAAGATGTTGCTCTAAAGATCTCGCA 94074

Qy 741 AATGCGATTTCTACAGTCCCGATTTATTTCATCATCATCAGATGATCATTCGCCGAG 800
Db 94073 AATGGAATTTCACTGTTCTTATTCATTCATTCATTCATTCATTCATTCATTCATTC 94014

Qy 801 TTATTGTCATCTTGATCATCTTTCTCGAATTTGGAATGTTATTTGATGATCATCTCAAAAC 860
Db 94013 CTACTACATCTTGCTTCACTTTCAGGGATCCCTAGATGTAAGTGTTTTTTTTTTTTAC 93954

Qy 861 TCACTGTAATTTGCCCTATTTCTGAAC 887
Db 93953 AACTTTCTTTTCTGCTGTTAAGATC 93927

RESULT 15
F7G19
LOCUS F7G19 110514 bp DNA linear 21N 16-SEP-1997
DEFINITION Sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1,
complete sequence.
ACCESSION AC000106
VERSION AC000106.1 GI:2342673
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 110514)
Osborne, B.I., Vysotskaia, V.S., Toriumi, M., Yu, G., Oji, O.,
Buehler, E., Conway, A.B., Conway, A.R., Dewar, K., Feng, J., Kim, C.,
Kurtz, D., Li, Y., Shinn, P., Sun, H., Davis, R.W., Ecker, J.R.,
Fedoripiel, N.A. and Theologis, A.
The sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1
Unpublished (1997)
2 (bases 1 to 110514)
Theologis, A.
Direct Submission
Submitted (20-JAN-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
3 (bases 1 to 110514)
Theologis, A.
Direct Submission
Submitted (17-FEB-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA

```



```

REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (27-FEB-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
5 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (03-MAR-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
6 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (02-APR-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
7 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (16-APR-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
8 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (26-AUG-1997) Plant Gene Expression Center, 800 Buchanan
St., Albany, CA 94710, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
9 (bases 1 to 110514)
Theologis,A.
Direct Submission
Submitted (16-SEP-1997) Plant Gene Expression Center, 800 Buchanan
St., Albany, CA 94710, USA
COMMENT
On Aug 26, 1997 this sequence version replaced gi:1541948.
The sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1.
FEATURES
Source
1. .110514
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F7G19"
971..2850
/gene="F7G19.1"
Join(971..1279,1540..1641,1688..1837,2134..2230,
2556..2665,2740..2850)
/gene="F7G19.1"
Note gb|N65461 comes from this gene.
EST gb|N65461 comes from this gene.
/codon_start=1
/evidence=not_experimental
/protein_id="AA570396.1"
/db_xref="GI:2342674"
/translation="MASCLQASNNLLPRSSSPSPHPPLSSSGRRNLKTPRYAFRA
KASAKTPMPPIPKOPPELSTLASIAANSPEKLNRPVNADVPYLDIPDSQOLMSSPA
QVRSVAYNEHPPTPPPLPSYLLDGRIVYIGMPYDDIFCFQIVPAVELVVAELM
YIQMLDPKPIYIINSTGTRDGETVHCVCVGAAGQACLLSAGTKGRKMPHHA
KEESEFCYFIFLIPWQLQVITNRDILVELLSKHTGNSVETVANYMRRPYNDAPKA
KFGVTRVSLGVIP"
complement(3700..3771)
/product="tRNA-Gly"
/note="codon recognized: GAA"
3942..6513
/gene="F7G19.2"
Join(3942..4046,4391..4507,4747..4854,4973..5164,
5263..5454,5551..5669,5747..5834,5916..6035,6216..6356,
6436..6513)
/gene="F7G19.2"
/note="Similar to probable Mg-dependent ATPase
(pir|S56671). ESTs gb|T46782.gb|AA04738 come from this
gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AA570397.1"
/db_xref="GI:2342675"

/translation="MAEDTSFEGQLASMTTDDIGRASRLIANBIRILKERSQRTNLD
LESVKEIKENQKIKLQKOLPYLVGNVILEMSPEDDAEDGANDLDSQKQKCV
VLTSTQTIPLPVGLVADSDTLKPGDLGVNKDSYLIDTLPSYDSRVKAMVDEK
PTDYNDIGLEKIQELVEALVLPMTHKQPEKLGIRPKGHVLLPGPGTKMILMAR
ACAAQNTATKLKAGPOLVQXFGIDGAKLVDAFLAKRSPCLIFIDEIDATGTRF
DSVSGDREVORTMLELNLQDGFSDDRDKVIAATNRADILDPALMRSGRLDKRKEF
PHPTTEARGILQIHSGKMKVNADVNFELARSTDDDFNGAQLKAVCEAGMALRERDA
TEVNDENEGIIQVQAKKASLNYYA"
complement(6693..9149)
/gene="F7G19.3"
Join(6693..6802,6892..7391,7471..7544,
7619..7721,7805..7920,8032..8127,8253..8639,8799..9149)
/gene="F7G19.3"
/note="Strong similarity to Oryza NADPH oxidase
(gb|X93301)."
/codon_start=1
/evidence=not_experimental
/protein_id="AA570398.1"
/db_xref="GI:2342676"
/translation="MLSQKLPTKDRNPVKRFAMNISYFFLENNKRIWLTWISICI
TLFTWFLQYKRTKTVFVVMCYCVTVKAGSAETLKFNNALILLPVCRNTITWLRKSL
IGSVPEDDNINPHKVAEGIAVGIGLHAISHLACDFPRLHAKNVEEPMKKEFGE
RPNYGFWMAGTDGWTGVTVLWLVAVYLAQWFRNRANLPKSLRGLTGFNAPWYS
HELFEVIVYLLIVHGVYVLSKWEYHTKTMVLAVPVLYAFELIEAFRPGAKAVY
LKVAVPGNVLISLYMSKPGKFKYTSQGYIYINCSDVPLQWHPFIIISASGDDYLSVH
IRLGDWTSQKSLYSKVCOLPSTQSGLFIADIGQANNITRPPRLIDGPGYAPAO
YRNYDVLLVGLGIGATPLSIIRDVLNNIKNONSIERGTNOHIKNYVATKRAYEVW
TRQGSLEWSEFSEVMNEVAYDSGMIEHLNYCTSVYEGDASALITMLOSLLHAKSG
IDVSGTRVTRFHPARNRWSRFAHVAVNVQNVGVFCGNTCIIGLKLKLAQPSRK
TTTKFHFHKEF"
complement(9310..10430)
/gene="F7G19.4"
Join(9310..9492,9582..9630,9745..9904,
10061..10430)
/gene="F7G19.4"
/codon_start=1
/evidence=not_experimental
/product="F7G19.4"
/protein_id="AA570399.1"
/db_xref="GI:2342677"
/translation="MESSSGETNKISRCATGSDNDPDEYVEITLVRDETINTMKA
KATRSVLISGLTKMVKUSFSARSLDPSKPGAFALGRFLAKNDAGRGWDEVA
KATPDLKAVEGKPKSGHCHGVSESEFVNFELFALVRRTGTTSSITKTELFPE
QITGNSFDRLQIPIFDWVKNLDRITGDEVKEIILASASANKLSKIENWDEVAALI
MEELDNLGYIEVRCLLSNGYTCULTSLKTLNP"
11553..13929
/gene="F7G19.5"
Join(11553..11812,11908..12119,12289..13011,13383..13344,
13419..13929)
/gene="F7G19.5"
/note="Similar to Arabidopsis luminal binding protein
(gb|D89342)."
/codon_start=1
/evidence=not_experimental
/protein_id="AA570400.1"
/db_xref="GI:2342678"
/translation="MSSLAIEGEQKLTGVIDLGTYSYSCGVYHKKHVEIANDQG
NRITPSWAPTDTERLIGBAKNAQAKNPRTIFDPKRLIGRKFPDPPDVORDIKFLPY
KVUNKDKPYIOVKVKEEKLSPETISAMILTMKEABAFGLKKIKDAVITVPAYE
NDAQOATKDAAGALAGNVVRIINEPTGAIAGLKKGGESNLLVYDGGGTFDVI
LTIDNGVFEVLSGTGDLGGDFDHRVMDYFIKLVKKYKNDIKSDHKALKLRRC
ELAKLSNQHQVRVIEISLFDGVSFELTRARFELNMDLFKTMEPVYKALKDAG
LKKSIDETVLVGGSTRIPKVQMLKDFDGPSPKTNPDFAVYGAAYQVGVLSGE
CGEYOKNLTLKNTDILLDVAPLSLGIIETVGGVNTIIPRNTVITPKTSQVFTYD
QQTVTIYVEGERSMTKDNRELKGFOLGILPAPRQVYRGVQLEVFEDVANGIL
QVAKDKVAKTSQITITNDKGLTEEEHEMIREAREFAEDKIKWEKIDARKMLT
YVNMKSTVADKELAKISDEKDKRGVGLKEALEWLENVNAEKEDYDEKLEVEL
VCDPVISYVETEGENEDTMEITWSYRIV"
complement(14446..15420)
/gene="F7G19.6"
Join(14446..15420)
/note="Similar to Glycine SRC2 (gb|AB000130). ESTs

```



```

gb|H75869.gb|727700.gb|ATTS089 come from this gene."
/codon_start=1
/evidence=not experimental
/protein_id="AAB70401.1"
/db_xref="GI:1922937"
/translation="MECRSLDLTIISAEDLKVOL-CKQDLVAVVISINGDARTKQTK
VDKDCGTFKWKQMKLTVDAAARDNRLTLVEIVADRRPIAGDKPVGEVSPVKELL
DQNGDEKTVTVAVRPNKAGSLKFSFKGSKITYGSSSGPHAPVFSAMDKTMD
QPVTAYPFGAPASAYAPGAPSSGYPQGHDDKHGVYGPQQAGYPAGTGGYPPP
GKYPQGGYGPYPQGGYGPYPQGGYGPYPQGGYGPYPQGGYGPYPQGGYGPYPQ
HGKAGAGMGLGLGLGLGLGLGLGLGLGLGLGLGLGLGLGLGLGLGLGLGLGL
17933..21752
/gene="F7G19.7"
join(17933..18538,18639..18781,18879..19029,19082..19181,
19238..19367,19470..20115,20194..20523,20583..20618,
20661..21032,21125..21194,21284..21369,21470..21577,
21678..21752)
/gene="F7G19.7"
/cote="Similar to Vicia sativa ENBP1 (gb|X95995)."
/codon_start=1
/evidence=not experimental
/protein_id="AAB70402.1"
/db_xref="GI:2342679"
/translation="MQVNFDETCDVIRMANEOTRSANGIGNGERSIPGIPDRLRC
KRSDKQRCCTMSMDKTCVCEKHVICAKKKAANSFRANOKKAKRRSSILGETDTYSE
GKDDDFELPVTSDIHYNGLASAKSNGRLEKRNKSLMRYSPETPMRSPSPRAVD

```

```

Query Match      4.9%; Score 77.2; DB 8; Length 110514;
Best Local Similarity 64.6%; Pred. No. 5.8e-10;
Matches 115; Conservative C; Mismatches 63; Indels 0; Gaps 0;

QY      836 ATGTATTGTAGACTTCAAACTCAACTGCTAATTGGCTATTCTGACCAACCAAGT 895
Db      41746 AGGTATATGCAGATACITTAGACATTCGTCTAGCTCTTTACCTATCCTTCAGCAACCCAT 41805

QY      898 GTGCTCCATCCGGTGGGTTCCTGGGTTCCGAAATTCGTGATCTGAACAGTCGTCATT 955
Db      41806 TTGTTCAATCCCTGGGTAGCTGGGTCCCTAGATCGAGATCAAGTAGCAACCTCT 41865

QY      956 GGCTATGTGGAAGCTAAGCTCACTAGCTCTGCCCTTAACATGTTGGTCAAGCT 1013
Db      41866 CGCTACATTGAGACCTCAAGCTCTTTGGGTTCTGGCTGCTGCAATTATTAGTCAAGGT 41923

```

Search completed: July 9, 2004, 07:01:42
Job time : 6290 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 01:16:27 ; Search time 714 Seconds
(without alignments)
9376.975 Million cell updates/sec

Title: US-09-857-525C-1
Perfect score: 1576
Sequence: 1 gcacgaggaacatttcgcg.....tcaaaaaaaaaaaaaaaa 1576

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002s.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1576	100.0	1576	3	AAA51347
2	987.4	62.7	2149	3	AAA51348
3	979.4	62.1	1632	3	AAA51353
4	594.2	37.7	2538	3	AAA51351
5	547.2	34.7	2160	3	AAA51350
6	529	33.6	538	3	AAA51354
7	266.2	16.9	282	6	ABL70761
8	259	16.4	514	3	AAA51357
9	148.4	9.4	702	3	AAA51349
10	139	8.8	178	6	ABL75404
11	95	6.0	542	3	AAA51355
12	50.6	3.2	1905	6	AB551409
13	50.6	3.2	3261	4	ABL10857
14	50.6	3.2	3951	4	ABL10856
15	49.6	3.1	969	7	ACG60984
16	46.8	3.0	472	8	ACH36558
17	43.8	2.8	498	3	AAA51356
18	43.8	2.8	1266	3	AAA51352
19	43	2.7	2000	7	ADA71938
20	42.5	2.7	324	7	ABT22888
21	42.5	2.7	1576	2	AA78927
22	42.4	2.7	1578	6	AA48616
23	42.4	2.7	1578	6	ABK84324

24	42.4	2.7	1578	7	ACA56515	Aca56515 Human sig
25	42.4	2.7	1691	3	ACA98774	Aac98774 Human pan
26	42.4	2.7	1774	6	ABU39755	Abi39755 Human NS
27	42	2.7	735	4	AAH07561	Aah07561 Human cDN
28	42	2.7	1167	6	AA48614	Aal48614 Human ins
29	42	2.7	1435	2	AAH06882	Aax06882 Disease a
30	42	2.7	2000	7	ADA71938	Ada71938 Rice gene
31	42	2.7	2223	4	AAH14839	Aah14839 Human cDN
32	42	2.7	2223	6	AA48615	Aal48615 Human ins
33	41.4	2.6	447	9	ADB57780	Adb57780 Toxicity-
34	41.4	2.6	990	6	ABQ69317	Abq69317 listeria
35	41.4	2.6	996	6	ABQ67749	Abq67749 listeria
36	41.4	2.6	81905	6	ABQ69244	Abq69244 listeria
37	41.4	2.6	82689	6	ABQ67198	Abq67198 listeria
38	41	2.6	491	9	ADE81890	Ad81890 Arabidops
39	41	2.6	2804	6	ABK35928	Abk35928 cDNA sequ
40	40.2	2.6	670	4	AA526003	Aas26003 Human cDN
41	40.2	2.6	670	7	ABX73344	Abx73344 Human nov
42	40	2.5	375	7	ABX49849	Abx49849 Bovine ES
43	39.6	2.5	843	5	ABV16074	Abv16074 Human pro
44	39.4	2.5	459	6	ABL93683	Abi93683 Arabidops
45	39.4	2.5	594	3	AA79823	Aac79823 Human sec

ALIGNMENTS

RESULT 1
AAA51347
ID AAA51347 standard; cDNA; 1576 BP.

AC AAA51347;

DT 26-SEP-2000 (first entry)

DE Z. mays sucrose non-fermenting 4 protein cDNA from clone csk1c.pk001.c15.

XX Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;

XX carbon catabolite repression; development; nitrogen partitioning; ss.

XX Zea mays.

XX Key Location/Qualifiers

FT 3. .1358

FT /*tag= a

FT /label= SNF4

FT /product= "Sucrose non-fermenting_4_protein"

XX partial

XX WO200036116-A2.

XX 22-JUN-2000.

XX 15-DEC-1999; 99WO-US029825.

XX 16-DEC-1998; 98US-0112564P.

XX (DUJO) DU PONT DE NEMOURS & CO E. I.

XX Allen SM, Heppard EP, Miao G, Weng Z;

XX WPI; 2000-431594/37.

XX P-PSDB; AAY96783.

XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins

XX involved in carbon catabolite repression in plants and seeds; useful for

XX controlling carbon and nitrogen partitioning pathways during plant growth

XX and development.

XX Claim 2; Page 31; 48pp; English.

XX This cDNA encodes maize (Zea mays) sucrose non-fermenting 4 (SNF4)

XX protein which is involved in carbon catabolite repression in plants and

CC

CC seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae, from clone
 CC cskic.pk001.c15 prepared from corn unpollinated developing silk 24 hours
 CC after emergence. The polynucleotides are used in plants to control carbon
 CC and nitrogen partitioning pathways during plant growth and development.
 CC The catabolite repression proteins would facilitate studies for better
 CC understanding the mechanism of catabolite repression in plants and could
 CC provide genetic tools to enhance or otherwise alter the accumulation of
 CC carbohydrates, lipids and proteins during plant growth and development
 XX
 SQ Sequence 1576 BP; 418 A; 336 C; 362 G; 460 T; 0 U; 0 Other;

Query Match 100.0%; Score 1576; DB 3; Length 1576;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACGAGGGAACATTTGCCGATGTCCTCTATCGAAGGCTGCCCACTGTATTTTCAGGCTA 60
 Db 1 GCACGAGGGAACATTTGCCGATGTCCTCTATCGAAGGCTGCCCACTGTATTTTCAGGCTA 60

Qy 61 TTTCGAGTCTGTCTCTCGGAGTTACAGAGTACAAATTTCTTTGAGCGGGGAATGGCGGC 120
 Db 61 TTTCGAGTCTGTCTCTCGGAGTTACAGAGTACAAATTTCTTTGAGCGGGGAATGGCGGC 120

Qy 121 ATGATGAGCGTCAACCTTACCATATCTGGGAGTTTGGCATAGTTAAACACACTTTACTTGA 180
 Db 121 ATGATGAGCGTCAACCTTACCATATCTGGGAGTTTGGCATAGTTAAACACACTTTACTTGA 180

Qy 181 CAAGGGAATATAACCAATAAACCCTTATCAAGTCCAAAGCACACCTGGAGCAGGATGA 240
 Db 181 CAAGGGAATATAACCAATAAACCCTTATCAAGTCCAAAGCACACCTGGAGCAGGATGA 240

Qy 241 ACATGAGTGTGTAATGAAATTTTCAACTGAGTGTCTGCTGAGTGGCAGCCGTTT 300
 Db 241 ACATGAGTGTGTAATGAAATTTTCAACTGAGTGTCTGCTGAGTGGCAGCCGTTT 300

Qy 301 CAGAGGACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGTGTGCTGTCTTG 360
 Db 301 CAGAGGACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGTGTGCTGTCTTG 360

Qy 361 AATATCTGAATTTGCATACATGCTATGATTTACTCCAGATCTGGCAAGGTTATTTGCC 420
 Db 361 AATATCTGAATTTGCATACATGCTATGATTTACTCCAGATCTGGCAAGGTTATTTGCC 420

Qy 421 TAGACATTAATTTACCTGTGAAGCAATCACTTCCATATTTCTCATGTAAGAGGGAATCTG 480
 Db 421 TAGACATTAATTTACCTGTGAAGCAATCACTTCCATATTTCTCATGTAAGAGGGAATCTG 480

Qy 481 TAGCTCTCTCTGGGACTCATTCAGAGTCAATTTGTTGCTCTCTTAGCCCATTTGGATT 540
 Db 481 TAGCTCTCTCTGGGACTCATTCAGAGTCAATTTGTTGCTCTCTTAGCCCATTTGGATT 540

Qy 541 TCATATCTATTTGCGGAGCTAGAACTCATGCTCGAATTCAGAGGAGCAGCTTG 600
 Db 541 TCATATCTATTTGCGGAGCTAGAACTCATGCTCGAATTCAGAGGAGCAGCTTG 600

Qy 601 AAACACACATATATCTGATGGAAGAGGCTAAGCGGCAAACTAATCGAAGAATGATA 660
 Db 601 AAACACACATATATCTGATGGAAGAGGCTAAGCGGCAAACTAATCGAAGAATGATA 660

Qy 661 GTGAGTGGCGGACCGCAACAGATCTAGTGCATGCGCCACCCCTTATGAGTCTTTGAGGACA 720
 Db 661 GTGAGTGGCGGACCGCAACAGATCTAGTGCATGCGCCACCCCTTATGAGTCTTTGAGGACA 720

Qy 721 TTGAGTAAAGCTTTTGAAGTGCATTTCTACAGTCCAGATTTATTTATCATCATCAT 780
 Db 721 TTGAGTAAAGCTTTTGAAGTGCATTTCTACAGTCCAGATTTATTTATCATCATCAT 780

Qy 781 CAGATGGAATCATTTCCCGAGTTATTTGATCTTGCATCTTCTGGAATTTTGAATGTA 840
 Db 781 CAGATGGAATCATTTCCCGAGTTATTTGATCTTGCATCTTCTGGAATTTTGAATGTA 840

Qy 841 TTGTGAGATACTTCAAAAACCTCAACTGGTAATTTGCTTATTTCTGAAACCAACGAGTGTGT 900

Db 841 TTGTGAGATACTTCAAAAACCTCAACTGGTAATTTGCTTATTTCTGAAACCAACGAGTGTGT 900

Qy 901 CCATTCGCTGGGTTCTGGGTTCCGAAATTTGGTATCTGACAGTCTGTCATTTGGCTA 960
 Db 901 CCATTCGCTGGGTTCTGGGTTCCGAAATTTGGTATCTGACAGTCTGTCATTTGGCTA 960

Qy 961 TGTTCGACCTAATGCTCTCACTTAGCTCTGCCCTTAACATGTTGTTCAAGCTCGAGTAA 1020
 Db 961 TGTTCGACCTAATGCTCTCACTTAGCTCTGCCCTTAACATGTTGTTCAAGCTCGAGTAA 1020

Qy 1021 GCTCAATACCAATTTGTGGATGACAAACGACTCCCTGCTTGAACCTTACTCTAGAAAGTGA 1080
 Db 1021 GCTCAATACCAATTTGTGGATGACAAACGACTCCCTGCTTGAACCTTACTCTAGAAAGTGA 1080

Qy 1081 TCACAGCCCTAGCAAAAGACAAAGTCTACACATGTTGCTGGATGAGATGACCAATTC 1140
 Db 1081 TCACAGCCCTAGCAAAAGACAAAGTCTACACATGTTGCTGGATGAGATGACCAATTC 1140

Qy 1141 ACCAGGCTTTGACAGCTTGGACCAAGATGCAATACACCTTTTGGATTTCTTAAACGCCAGA 1200
 Db 1141 ACCAGGCTTTGACAGCTTGGACCAAGATGCAATACACCTTTTGGATTTCTTAAACGCCAGA 1200

Qy 1201 GATTCAGATGTCCTCCGCTCTGATCTTTGCTGAAGTGTATGGAGCGACTGGCTAATC 1260
 Db 1201 GATTCAGATGTCCTCCGCTCTGATCTTTGCTGAAGTGTATGGAGCGACTGGCTAATC 1260

Qy 1261 CTGGGCTGGCGCGGCTGTCATTTGGAAGCTGGGACCAACCTGTTGGAGGCGCATCATAT 1320
 Db 1261 CTGGGCTGGCGCGGCTGTCATTTGGAAGCTGGGACCAACCTGTTGGAGGCGCATCATAT 1320

Qy 1321 CACTAAGTGCATTTTCAAGTTCTTGTGAGCTTTGTGAGGAGAGAGGAGCATGGCGAGT 1380
 Db 1321 CACTAAGTGCATTTTCAAGTTCTTGTGAGCTTTGTGAGGAGAGAGGAGCATGGCGAGT 1380

Qy 1381 TTTTAAACAGTAGAACCCCTCTGGGCTTTGGGCTCTTGGGATTAATTTGCTATCTGTC 1440
 Db 1381 TTTTAAACAGTAGAACCCCTCTGGGCTTTGGGCTCTTGGGATTAATTTGCTATCTGTC 1440

Qy 1441 AGGCACTCNGGTTCTGATGCTGTAAGTCTCTTTTGTGTTTGTGTTTACCATTT 1500
 Db 1441 AGGCACTCNGGTTCTGATGCTGTAAGTCTCTTTTGTGTTTGTGTTTACCATTT 1500

Qy 1501 GGCCTCCCTGTTTAAACAAACATACATCAGTGTGATTTTATCTATTGATTTGTTCAA 1560
 Db 1501 GGCCTCCCTGTTTAAACAAACATACATCAGTGTGATTTTATCTATTGATTTGTTCAA 1560

Qy 1561 AAAAAAAAAAAAAA 1576
 Db 1561 AAAAAAAAAAAAAA 1576

RESULT 2

AAAS1348
 ID AAAS1348 standard; cDNA; 2149 BP.
 XX
 AC AAAS1348;
 XX
 DT 26-SEP-2000 (first entry)
 XX
 DE Rice sucrose non-fermenting 4 protein cDNA from clone r10n.pk083.i18.
 XX
 KW Rice; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 carbon catabolite repression; development; nitrogen partitioning; ss.
 XX
 OS Oryza sativa.
 XX
 FH Key
 CDS
 FT 185..1666
 FT /*tag= a
 FT /product= "Sucrose_non-fermenting_4 protein"
 XX
 PN WO200036116-A2.

```

XX PD 22-JUN-2000.
XX PF 15-DEC-1999; 99WO-US029825.
XX PR 16-DEC-1998; 98US-0112564P.
XX PA (DUPO) DU PONT DE NEMOURS & CO E. I.
XX PI Allen SM, Heppard EP, Xiao G, Weng Z;
XX DR WPI; 2000-431594/37.
XX DR P-PSDB; AAY96784.
XX PT New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
XX PT involved in carbon catabolite repression in plants and seeds, useful for
XX PT controlling carbon and nitrogen partitioning pathways during plant growth
XX PT and development.
XX PS Claim 2; Page 33-34; 48pp; English.
XX CC This cDNA encodes rice (Oryza sativa) sucrose non-fermenting 4 (SNF4)
XX CC protein which is involved in carbon catabolite repression in plants and
XX CC seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
XX CC Arabidopsis thaliana and Saccharomyces cerevisiae, from clone
XX CC r10n.pk083.118 prepared from rice 15 day old leaf. The polynucleotides
XX CC are used in plants to control carbon and nitrogen partitioning pathways
XX CC during plant growth and development. The catabolite repression proteins
XX CC would facilitate studies for better understanding the mechanism of
XX CC catabolite repression in plants and could provide genetic tools to
XX CC enhance or otherwise alter the accumulation of carbohydrates, lipids and
XX CC proteins during plant growth and development
XX SQ Sequence 2:49 BP; 523 A; 492 C; 501 G; 633 T; 0 U; 0 Other;

Query Match 62.7%; Score 987.4; DB 3; Length 2149;
Best Local Similarity 83.2%; Pred. No. 6.9e-266;
Matches 1123; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 5 GAGGGAACATTTCGGATGTCCTATCGAAGGCTGCCACATGTATTTCAGGCTATTG 64
DB 316 GACAGAACATTTCGCAATGTCCTGTTGAAGTTGTCCACCCTTTTCGCGCATATG 375
QY 65 CAGTCTGTCTCTCGGATTCACAGATGACAAATCTTTGTGGACGGGAATGGCGCATGA 124
DB 376 CAGTTTGTCCCTGGGATTCATCAGTACAAATTTTGTGGATGGGAATGGCGCACGA 435
QY 125 TGAGGCTAACCTACCATATCTGSGAGTTTGGCAATGATTAACACACTTTACTTGCACAG 184
DB 436 TGAGCGGCAACCTACCATTAACAGGAGACTATGTTGTGTTATACACTTTATGCTTGACTAG 495
QY 185 GGAATATAACCAATAAACCCTTATCAAGTCCAAAGCACACCTTGGAGAGGATGAACAT 244
DB 496 GGACTTTGACAGTAATTAACAAATTTGAGCCCTAGTACACCTTGGAGTAGATGAACAT 555
QY 245 GGATGTGGATAATGAAATTTTCAACGTAAGGTTACGTTGTGAGTAGGACCGTTTCAGA 304
DB 556 GGATGTGGACACGACAAATTTTCAACGTAAGTCTGTTCTTCTGCTGATGGCATTTATCAGGA 615
QY 305 AGGTACTCTGAGATTTTCAGAGGCTGCAATACAAATATCTAGTGTGCTGTTCTGATATA 364
DB 616 AGGTCTCTGAGAAATTTTCAGAGGAGCTATACAGATCTCTAGTGTGCTGTTAGCGGATTT 675
QY 365 TCTGAATTTGCATACATGCTATGATTTTACTCCAGATTTCTGGCAAGTTATTCCTCCATGA 424
DB 676 TCTGAATGGCAAACTGGGTATGATTTTACTCCAGATTTCTGGCAAGTATCTGCTCTAGA 735
QY 425 CATTAATTTACTGTGGAAGCAATCAITTCATATTTCTCCATGAAACAGGGGATTCCTGTAGC 484
DB 736 GGTAAATTTGCTGTGGAAGCAATCTTTTCATATTTCTCATGAAACAGGGAATTCCTGTGGC 795
QY 485 TCTCTCTGGACTCATTCAGAGGTCATTTGTTGCTCTCTCTCTAGGCCATGGAATTCAT 544

```

RESULT 3
 AA51353
 ID AAA51353 standard; cDNA; 1632 BP.
 XX
 AC AAA51353;
 XX
 DT 26-SEP-2000 (first entry)
 XX
 DE Wheat sucrose non-fermenting 4 protein cDNA from clone wll.pk0002.b3.
 XX
 KW Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 XX carbon catabolite repression; development; nitrogen partitioning; ss.
 OS *Triticum aestivum*.

XX Key Location/Qualifiers
 FH 10.1393
 FT /*tag= a
 FT /product= "Sucrose_non-fermenting_4_protein"
 XX WO200036116-A2.
 PN 22-JUN-2000.
 XX
 PD 15-DEC-1999; 99WO-US029825.
 PF 16-DEC-1998; 98JS-0112564P.
 XX
 XX {DUPO } EU PONT DE NEMOURS & CO S I.
 XX Allen SM, Heppard EP, Miao G, Weng Z;
 XX WPI; 2000-431594/37.
 DR P-PSDB; AAY96789.
 XX
 XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 XX
 XX Claim 18; Page 43; 48pp; English.
 PS
 XX This cDNA encodes wheat sucrose non-fermenting 4 (SNF4) protein which is
 CC involved in carbon catabolite repression in plants and seeds. The cDNA
 CC was isolated, based on similarity to SNF4 proteins from Arabidopsis
 CC thaliana and Saccharomyces cerevisiae, from library will prepared from
 CC wheat leaf from 7 day old seedling. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development
 XX
 SQ Sequence 1632 BP; 431 A; 370 C; 389 G; 442 T; 0 U; 0 Other;

Query Match 62.1%; Score 979.4; DB 3; Length 1632;
 Best Local Similarity 83.5%; Pred. No. 1.1e-263;
 Matches 1125; Conservative 3; Mismatches 216; Indels 6; Gaps 1;
 12 CATTTCGGATGCTCTCTATCGAGGCTGCCCACTGTATTCAGGCTATTTCAGTCTG 71
 2 CACGAGCCGATGCTCCAGTCGAAGTTCGCTTCAAGTTCGCTTCAAGTTCGCAACTTA 61
 72 TCTCTGGGATTCAGAGTACAAATTCCTTTGGACGGGGAATGGCGCATGATGAGCGT 131
 62 CTTCCAGGATTTATCAGTACAAATTCCTTTGGACGGGGAATGGCGCATGATGAGCGG 121
 132 CACCTACCATATCTGGGGAGTTGGCATAGTTAAACACTTTTACCAAGGGAATAT 191
 122 CACCTACTACTAATCTGGAGAGTATGGGGTGGTAAACACTTTTATATCTTCAAGGGAAT 181
 192 AACCAATAAACACTTATCAAGTCCAGGACACCTCGAGGACGATGACATCGATG 251
 182 GACCACATAATCTACTGACGACCCCACTACCTGGGAGCAGG-----ATGGATGTG 235
 252 GATAATGAAATTTCAACGTTACGTTGTCAGATGSCACCGTTTCAGAGGTACT 311
 236 GACAGTACAGTTTTCACGAAATGGGTTGTTGTCGATGTTGCCCTTCAGGAGGTCT 295
 312 CTGAGAGTTTCAGAGGCTCAATACAAATATCTAGTGTGTTCTGTAATATCTGAAT 371
 296 CCAAGAAATCTCAGAGGCTGTATACAGATCTCTAGTGTGTTGTTGCTGAGTATCTGAAT 355
 372 TCGATACATGCTATGATTTCCTCCAGATCTCTGGCAGGTTATTCCTAGCATTAAT 431
 356 GCGCATACAGGCTATGACCTTACTACAGATCTCTGGAAAGGTCAATTGCTCTGGACATTAAT 415

QY 432 TTACCTGTGAAGCAATCAATTCATATTTCTCCATGAACAGGGGATTCCTGTAGCTCTCTC 491
 DB 416 TTACCTGTGAAGCAATTCCTTCATATTTCTCCATGAACAGGGGATTCCTGTAGCTCTCTC 475
 QY 492 TGGGACTCATTGAGAGTCAATTTGTTGTTCTCTCTAGCCCATTTGATTTTCACTCAT 551
 DB 476 TGGGATTCATTGAGGGGTGAGTTTGTGGCTTCTGAGCCCACTGGATTTTATCTTATA 535
 QY 552 TTGGCGGAGCTAGAACTCATGCTCGAACTTACACAGAGAGGAGCTTGAACACACACT 611
 DB 536 TTGAGAGAGCTGAACTCATGCTCAAACTTACAGAGAGGAGAGCTTGAACACACACT 595
 QY 612 ATATCTGATGGAAGAGGCTTAAGCGGCAAACTTAATGGAAGAAATGATAGTCACTAGTGGCG 671
 DB 596 ATATCTGCTGGAAGAGGCTTAAGCGGCAAACTTAATGGAAGAAATGATAGTCACTAGTGGCG 655
 QY 672 CCGCAACAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTTGAAGGACATTTGAGTAAAG 731
 DB 656 TCAATCAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTTGAAGGAGTATTTGCCATGAA 715
 QY 732 CTTTTCGAAATGCGCATTTCTACAGTCCAGTTTATTTTATTCATCATCATCATCATCAT 791
 DB 716 ATACTGGAAGTGGCATTTCTACAGTCCCAATCTCTATTCATCTCATCATCATCATCAT 775
 QY 792 TTCCCGCAGTTATGCACTTGCATCTTCTGCAATTTTGAATATTTGAAATGATTTGTAGATAC 851
 DB 776 TTTCGGCAGCTTGCATCTTGCATCTTGCATCTTGCATCTTGCATCTTGCATCTTGCATCT 835
 QY 852 TTCAAAATCAACTGGTAAATTTGCTATTTGAAACCAACAGTGTGCTTCCATTCGCTG 911
 DB 836 TTCAAAATCAACTGGTAAATTTGCTATTTGAAACCAACAGTGTGCTTCCATTCGCTG 895
 QY 912 GGTTCCTGGGTTCCGAAATTTGCTATCTGAACAGTGTGCTTGTGCTATTTGTGACCT 971
 DB 896 GGTTCCTGGGTTCCGAAATTTGCTATCTGAACAGTGTGCTTGTGCTATTTGTGACCT 955
 QY 972 AATGCTCATTAGTCTTGCCTTAAACATGTTGTTCAAGCTGGAGTAAAGCTCAATACCA 1031
 DB 956 AATACATCTTGTAGCTTGCCTTAACTTGTGTTCAAGCTGGAGTAAAGCTCAATACCA 1015
 QY 1032 ATTGTGATGACAAAGCTTCCCTGCTTGCATCTTGCATCTTGCATCTTGCATCTTGCATCT 1091
 DB 1016 ATTGTGATGACAAAGCTTCCCTGCTTGCATCTTGCATCTTGCATCTTGCATCTTGCATCT 1075
 QY 1092 GCAAAAGCAAGGCTTACACATTTTGGCTGATGAGATGACCAATTCACAGGCTTTG 1151
 DB 1076 GCGAAGCAAGGCTTACACATTTTGGCTGATGAGATGACCAATTCACAGGCTTTG 1135
 QY 1152 CAGCTTGGACAGGATCCCAATACACCTTTTGGATTTTAAAGGCGAGATGCGAGATG 1212
 DB 1136 CAGCTCGGGCAAGAGCGGATTTCACTTTTGGATTTTAAAGGCGAGATGCGAGATG 1195
 QY 1212 TGGCTCGGCTGATCTCTTGTGCTGAGGATGAGGAGTGGCTTAACTCTGGGCTGGG 1271
 DB 1196 TGTCTCAGTCTGACCTTTGCTGAAAGTATGAGGATTTGGCTTAACTCTGGGCTGGG 1255
 QY 1272 CGGCTGTTTCTTGTGAGGCTGGGAGCAAAAGTGTGGAGGGGATCATATCATCTAAGTGAC 1331
 DB 1256 CGGCTGTTTCTTGTGAGGCTGGGAGCAAAAGTGTGGAGGGGATCATATCATCTAAGTGAC 1315
 QY 1332 ATTTTCAAGTTCTTGTGAGGCTTTGTA 1358
 DB 1316 ATATTCAGTTGCTGTGAGCTAGCGA 1342

RESULT 4
 AAA51351
 ID AAA51351 standard; cDNA; 2538 BP.
 XX
 AC AAA51351;
 XX
 DT 26-SEP-2000 (first entry)

```

XX Soybean sucrose non-fermenting 4 protein cDNA from clone sfl1.pk0004.b4.
DE
XX
XX
XX Soybean; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
KW carbon catabolite repression; development; nitrogen partitioning; ss.
XX
XX Glycine max.
OS
XX
XX
FH Key Location/Qualifiers
FT 395..1873
FT CDS
FT /tag= a
FT /product= "Sucrose non-fermenting_4_protein"
XX
XX WO200036116-A2.
XX
XX 22-JUN-2000.
XX
XX 15-DEC-1999; 99MO-US029825.
XX
XX 16-DEC-1998; 98US-0112564P.
XX
XX (DUF0 ) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Heppard EP, Miao G, Weng Z;
XX
XX WPI: 2000-431594/37.
XX P-PSDB; AAY96787.
XX
XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
XX involved in carbon catabolite repression in plants and seeds, useful for
XX controlling carbon and nitrogen partitioning pathways during plant growth
XX and development.
XX
XX Claim 2; Page 39-40; 48pp; English.
XX
XX This cDNA encodes soybean (Glycine max) sucrose non-fermenting 4 (SNF4)
XX protein which is involved in carbon catabolite repression in plants and
XX seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
XX Arabidopsis thaliana and Saccharomyces cerevisiae, from library sfl1
XX prepared from soybean immature flower. The polynucleotides are used in
XX plants to control carbon and nitrogen partitioning pathways during plant
XX growth and development. The catabolite repression proteins would
XX facilitate studies for better understanding the mechanism of catabolite
XX repression in plants and could provide genetic tools to enhance or
XX otherwise alter the accumulation of carbohydrates, lipids and proteins
XX during plant growth and development
XX
XX Sequence 2538 BP; 670 A; 463 C; 571 G; 834 T; 0 U; 0 Other;
XX
XX
XX Query Match 37.7%; Score 594.2; DB 3; Length 2538;
XX Best Local Similarity 66.8%; Pred. No. 1.2e-155;
XX Matches 896; Conservative 0; Mismatches 433; Indels 12; Gaps 3;
XX
XX 8 GGAACATTGGCGATGTCCTCTATCGAAGCTGCCCCACCTGTAATTCAGGCTATTGGCAG 67
XX 529 GGAGCTCTACCGATGTCGCCAGTGGAGAGTGTCCAACTGTGTTCAAGTGAATTAATA 588
XX
XX 66 TCTGTCTCTCGGATTCACGAGTACAAATCTTTGTGGAGCGGGAATGGCGCATGATGA 127
XX 589 CTGGCCACCGGTTACCATCAGTACAAAGTTTTTGTGTGAGGAATGGCGGATGATGA 648
XX
XX 128 GGCTCAACCTACCATATCTGGGAGTTTGGCATAGTTAAACACACTTTACTTGAAGGGA 187
XX 649 ACATCAACCTTATGTACCTGGAGATATGGATAGTCAACACTGTCTTATGGCCACTGA 708
XX
XX 188 ATATACCAAAATAAACCTTATCAAGTCCAGCACACCTGGAGCAGGATGAATGGA 247
XX 709 TCTTAACATACGCTGTTTACCTCCAGACGTTGCTCTGGAATAG-----CATGGA 762
XX
XX 248 TGTGGATAATGAAATTTTCAAGTACGGTTAGTTGTGAGTGGCACCGTTTCAGAAGG 307
XX 763 TGTGGATAATGAGTCTTTTGGCGGAATGGCCGGTTGACCGATGACTTTGAGTGAGGT 822

```

RESULT 5
 ID AAA51350 standard; cDNA; 2160 BP.
 AC AAA51350;
 DT 26-SEP-2000 (first entry)
 XX Soybean sucrose non-fermenting 4 protein cDNA from clone ses4d.pk0040.b1.
 XX
 XX Soybean; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 XX carbon catabolite repression; development; nitrogen partitioning; ss.
 XX Glycine max.
 XX
 XX Key Location/Qualifiers
 XX CDS 185..1633
 XX /*tag= a
 XX /product= "sucrose_non-fermenting_4_protein"
 XX
 XX WO200036116-A2.
 XX
 XX 22-JUN-2000.
 XX
 XX 15-DEC-1999; 99WO-US029825.
 XX
 XX 16-DEC-1998; 98US-0112564P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Allen SM, Heppard EP, Miao G, Weng Z;
 XX WPI; 2000-431594/37.
 XX P-ESDB; AAY96785.
 XX
 XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 XX involved in carbon catabolite repression in p-ants and seeds, useful for
 XX controlling carbon and nitrogen partitioning pathways during plant growth
 XX and development.
 XX
 XX Claim 2; Page 36-37; 48pp; English.
 XX
 XX This cDNA encodes soybean (Glycine max) sucrose non-fermenting 4 (SNF4)
 XX protein which is involved in carbon catabolite repression in plants and
 XX seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
 XX Arabidopsis thaliana and Saccharomyces cerevisiae, from library ses4d
 XX prepared from soybean embryogenic suspension 4 days after subculture. The
 XX polynucleotides are used in plants to control carbon and nitrogen
 XX partitioning pathways during plant growth and development. The catabolite
 XX repression proteins would facilitate studies for better understanding the
 XX mechanism of catabolite repression in plants and could provide genetic
 XX tools to enhance or otherwise alter the accumulation of carbohydrates,
 XX lipids and proteins during plant growth and development
 XX
 XX Sequence 2160 BP; 588 A; 392 C; 485 G; 693 T; 0 U; 2 Other;
 XX
 XX Query Match 34.7%; Score 547.2; DB 3; Length 2160;
 XX Best Local Similarity 66.0%; Pred. No. 1.6e-142;
 XX Matches 990; Conservative 0; Mismatches 413; Indels 45; Gaps 5;
 XX
 XX 3 GACATTTGGCGATGCTCTCTATGAGAGGCTGCGCCACTGTATTTTCAAGCTATTTCAGT 68
 XX 317 GAACTTCTACAAATGTGCGCTGTGAGAGGCTGCGCCACTGTGTTTCAAGTATTTCATAGC 376
 XX
 XX 69 CTCTCTCTGGGATTACAGAGTACAATCTTGTGAGCGGGGAATGGCGCATGTAG 128
 XX 377 TTGGTACCTGGTCATCATCATGATACAGATTTTGTGATGGAGATGGCGCATGATGAC 436
 XX
 XX 129 CGTCAACCTACCATATCTGGGAGTTTGGCATAGTTTAAACACACTTTTACCTTGACAAAGGAA 188
 XX 437 CTTCAACCTTGTGAATCTGGAGATATGGATGTTTAACTGTTTCACTGTTTCACTGGTACTGAT 496
 XX
 XX 189 TATAACCAATAAACACCTTATCAAGTCCAAAGCACACCTTGGAAAGCAGGATGAACATGGAT 248

Db 497 CCTAATATTTTACCTGTTTAACTCCAGACATAGTTCTGGG-----CTAATACGGAT 550
 QY 249 GTGGATAATGAAATTTTCAACGTACGGTTAGCTGTGTCAGATGGCACCGCTTTCAGAGGT 308
 Db 551 GTGCAACACGAGGCTTTTCGAGCATGTTCCGATTTGACAGATGATTAAGTAATGTG 610
 QY 309 ACTCTG---AGAGTTTCAGAGGCTGCAATACAAATATCTAGTGTGCTGTTCTGAATAT 365
 Db 611 TTAATGCAAGAATATCTGATTTGATATACAGACTCTCGTCAGCGTATTTCTGCAATTC 670
 QY 366 CTGAATTTGCTATACATGCTATGATTTACTCCAGATTTCTGCAAGTATTTCCCTAGAC 425
 Db 671 CTATCTATGAGTACAGCATGGAATTTCTCTGAGTCAGGCAAGTTGTTACCTTGGAT 730
 QY 426 ATTAATTTACCTGTGAAGCAATCATTTCCATATTTCTCCATGAACAGGGATTTCTGTAGCT 485
 Db 731 GTTGATCTGCTGTGAAACAGGCAATTTCCATATCTGATGACAGGGAATTTCCCATTTGCT 790
 QY 486 CCTCTCTGGGACTCATTTGAGAGGTTCAATTTGTTGTTCTCTCTCTAGCCCATTTGATTCATA 545
 Db 791 CCTCTCTGGGACATCTGCAAGGGGCGATTTGTTGAGTTCTTTAGTGCCTTGGATTTTAT 850
 QY 546 CTCATATTTGCGGAGCTAGAAACTCATGCTCGAACTTTGACAGAAGAGCAGCTTCAAAACA 605
 Db 851 TTAATTTTAAGAGAGCTCGGAATCATGGTCCATCTTAACAGAAGAGAGCTTGAACA 910
 QY 606 CACATATATCTGCATGGAAGAGGCTTAAGCGGCAAACTTAATGGAAGAAATGATAGTCAG 665
 Db 911 CATACCATATCAGCTTGAAGAGGAGAAATGAGCA----- 946
 QY 666 TGGGACCGCAACAGCATCTAGTGTGATGCCACCCCTTATGAGTCTTGAGGGACATTCGA 725
 Db 947 ---GGATTTACAAATGTTTATCGTGCGGGCCATACGATAATTTGAAAGAGATTCCT 1003
 QY 726 GTAAAGCTTTTGCAAATGCGCAATTTCTACAGTGCCAGTTATTTATTTATCATCATCAGAT 785
 Db 1004 GTGAAGATCTGCAACATGGAATTTCAACAGTTCTCTATTTATCTCAGA-----AGAT 1057
 QY 786 GGATCATTCGCGCAGTTATGTCATCTTGCATCACTTCTGGAATTTTGAATGATTTGT 845
 Db 1058 GGTTCATTTCCAGAGTACTACATCTTGCTTCACTTTTCAGGAATCTCTAAATGCAATTTGC 1117
 QY 846 AGATACCTTCAAAACTCACTGTTAAATTTGCTATTTGCAACCAACAGCTGCTCCATTT 905
 Db 1118 AGTATTTTGAAGATTTGTTCTAGTTCATGCTCTATCTTCAACTTCAATTTTGCAATTC 1177
 QY 906 CCGCTGGGTTTCTGGGTTCCGAAATTTGGTATCTGAACAGTCTGCTTCCATTTGGTATGTTG 965
 Db 1178 CTTGTGGGACGTTGGTGGCCAAATTTGGGAGTCAATCGCGGCTCTAGCAATGTTG 1237
 QY 966 CGACCTAATGCTCCTAGCTGCTGCTTAACTGTTGGTTCAAGCTGGAGTGAAGTCA 1025
 Db 1238 AGACCAAAATGCTTCACTTCTCAGCCCTAAACTTATTTAGTTCAAGTCAAGTAAAGTTCA 1297
 QY 1026 ATACCAATTTGGATGACACAGGCTCTCTGCTGACACTTACTCTAGAAGTGAATCACA 1085
 Db 1298 ATACCAATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1357
 QY 1086 GCCTTAGCAAAAGACAGGTTCTACACATGTTTGGCTGGATGAGATGACCATTTCCACAG 1145
 Db 1358 GCTTTGGCAAAAGACAGACTTATACCATTAATCTTGTGATGAATGATGATGATGATGATGAT 1417
 QY 1146 GCTTTGCGAGTTGGAAGATGCAATACACTTTTGGATTTCTTTAAGCCGAGATGATGC 1205
 Db 1418 GCATTTGCAATTTGGGCCAGGATTTCTTATAATTAATTTATG---AGCTGAGTTTCTCAAGATGT 1474
 QY 1206 CAGATGTCCTCGGCTCTGATCCTTTGCTGAAGGTGATGAGGACTTGGCTAATCTCTGGG 1265
 Db 1475 CAGATGTTTGGCAACTGATTTCTTGCATAAAGTGAAGAGCTTTGGCAAGTCCAGGT 1534
 QY 1266 GTGCGGGGGGTGTTCAATTTGGAAGCTGGAGGAGAAACGTTGTGGAGGCAATCATATCACTA 1325

Db 1535 GTGAGCGGCTTGTAATTTGGAGCTGGCAGCAGCGGCTAGAGGCAATCATAGCACTG 1594
 QY 1326 AGTGACATTTTCAAGTTCTTGCTGAGCT 1353
 Db 1595 AGTGACATATTCAACTTCTTCCITGGTT 1622

RESULT 6

AAA51354

ID AAA51354 standard; cDNA; 538 BP.

XX AC AAA51354;

DT 26-SEP-2000 (first entry)

XX Maize sucrose non-fermenting 4 protein cDNA from clove csk1c.pk001.cl5.

DE Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;

XX carbon catabolite repression; development; nitrogen partitioning; ss.

OS Zea mays.

XX Key Location/Qualifiers

FT CDS 347..523

FT /*tag= a

FT /product= "Partial_SNF4_protein"

FT 494..496

FT /*tag= b

FT /note= "Encodes Xaa which is not defined"

PN WO200036116-A2.

XX 22-JUN-2000.

XX 15-DEC-1999; 99WO-US023825.

XX 16-DEC-1998; 98US-0112564P.

XX (DUPO) DU FONT DE NEMOURS & CO E I.

XX Allen SM, Heppard EP, Miao G, Weng Z;

XX WPI; 2000-431594/37.

XX P-PSDB; AAY36789.

XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins

PT involved in carbon catabolite repression in plants and seeds; growth;

PT controlling carbon and nitrogen partitioning pathways during plant growth

PT and development.

XX Claim 18; Page 45; 48pp; English.

XX This partial cDNA encodes maize (Zea mays) sucrose non-fermenting 4

CC (SNF4) protein which is involved in carbon catabolite repression in

CC plants and seeds. The cDNA was isolated, based on similarity to SNF4

CC proteins from Arabidopsis thaliana and Saccharomyces cerevisiae, from

CC library csk1c prepared from corn unpollinated developing silk 24 hours

CC after emergence. The polynucleotides are used in plants to control carbon

CC and nitrogen partitioning pathways during plant growth and development.

CC The catabolite repression proteins would facilitate studies for better

CC understanding the mechanism of catabolite repression in plants and could

CC provide genetic tools to enhance or otherwise alter the accumulation of

CC carbohydrates, lipids and proteins during plant growth and development

XX Sequence 538 BP; 146 A; 115 C; 120 G; 156 T; 0 U; 1 Other;

XX Query Watch 33.6%; Score 525; DB 3; Length 538;

XX Best Local Similarity 98.9%; Pred. No. 1.1e-137;

XX Matches 532; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

XX 8 GGAACATTGGCGATCTCTCCATCGAAGCTGCCCCACTGTATTTCAGGCTATTGGAG 67

Db 1 GGAACATTGGCGATCTCTCCATCGAAGCTGCCCCACTGTATTTCAGGCTATTGGAG 60

QY 68 TCTGTCTCTCGGATTCACGAGTACAAATCTTTGTGACGGGGAATGGCGCATGATGA 127
 Db 61 TCTGTCTCTCGGATTCACGAGTACAAATCTTTGTGACGGGGAATGGCGCATGATGA 120
 QY 128 GCGTCAACCTTACCATAATCTGGGAGTTTGGCATAGTTAAACACACTTTCTTGAAGA 187
 Db 121 GCGTCAACCTTACCATAATCTGGGAGTTTGGCATAGTTAAACACACTTTCTTGAAGA 180
 QY 188 ATATAACCAATTAACACCTTATCACTCCCAAGCACAACCTGGGAAGCAGGATGAACATGGA 247
 Db 181 ATATAACCAATTAACACCTTATCACTCCCAAGCACAACCTGGGAAGCAGGATGAACATGGA 240
 QY 248 TGTGGATAATGAATAATTTTCAACGTTACGTTTACGTTTCTGATGCGCACCGTTTCAGAAGG 307
 Db 241 TGTGGATAATGAATAATTTTCAACGTTTACGTTTCTGATGCGCACCGTTTCAGAAGG 300
 QY 308 TACTCTGAGAGTTTCTGAGAGCTGCAATPACAAATATCTAGTGTCTGTTCTGAACTCT 367
 Db 301 TACTCTGAGAGTTTCTGAGAGCTGCAATPACAAATATCTAGTGTCTGTTCTGAACTCT 360
 QY 368 GAATTTGCATACATGCTATGATTTACTCCAGATTTCTGCAAGGTTATTCGCTTAGCAT 427
 Db 361 GAATTTGCATACATGCTATGATTTACTCCAGATTTCTGCAAGGTTATTCGCTTAGCAT 420
 QY 428 TAATTTTACCTGTGGAAGCAATCATTTCCATATTTCTCCATGAACAGGGGATTCCTGTAGCTCC 487
 Db 421 TAATTTTACCTGTGGAAGCAATCATTTCCATATTTCTCCATGAACAGGGGATTCCTGTAGCTCC 480
 QY 488 TCTCTGGGACTCAATTCAGAGTCAATTTGTTGTTGTTCTCTCTAGCCCATTTGGATTTCATA 545
 Db 481 TCTCTGGGACTCAATTCAGAGTCAATTTGTTGTTGTTCTCTCTAGCCCATTTGGATTTCATA 538

RESULT 7

ABL70761

ID ABL70761 standard; cDNA; 282 BP.

XX AC ABL70761;

XX 14-MAY-2002 (first entry)

DT Corn tassell-derived polynucleotide (cdps) SEQ ID NO:135.

XX Corn; corn tassell-derived polynucleotide; cdps; hybrid breeding; CDPs;

XX inheritance; characteristic; growth; development; disease resistance;

XX environmental adaptability; quality; yield; molecular marker;

XX multigene trait; plant breeding; corn tassell; gene; ss.

XX Zea mays.

OS US2001051335-A1.

XX 13-DEC-2001.

XX 16-APR-1999; 99US-00294093.

XX 21-APR-1998; 98US-0082567P.

XX (LALG/) LALGUDI R V.

XX (ITOL/) ITO L Y.

XX (SHER/) SHERMAN B K.

XX Lalgudi RV, Ito LY, Sherman BK;

XX WPI; 2002-163647/21.

XX Novel purified corn tassell-derived polynucleotide useful for determining

PT altered gene expression, to recover regulatory elements and to follow

PT inheritance of desirable characteristics through hybrid breeding

PT programs.

XX Claim 1; SEQ ID NO 135; 201pp; English.

XX The present sequence describes a purified corn tassel-derived
 CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
 CC selected from those given in ABL70627 to ABL76833. The cdps sequences
 CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
 CC can be used for determining altered gene expression, to recover
 CC regulatory elements and to follow inheritance of desirable
 CC characteristics through hybrid breeding programs. (I) are also useful in
 CC the evaluation, and alteration of desired characteristics associated with
 CC growth and development, disease resistance, environmental adaptability,
 CC quality and yield, and as molecular markers for studying inheritance of
 CC multi-gene traits in a plant breeding program. (I) can be used to produce
 CC a tassel-specific profile of gene transcription, a transcript image, to
 CC clone regulatory elements for use in transformation vectors, to express a
 CC polypeptide, to identify, isolate or extend identical or related corn
 CC tassel nucleic acid sequences from DNA libraries, in nucleic acid
 CC hybridisation or amplification technologies, as query sequences to
 CC determine homology of known sequences, as probe for use in Southern or
 CC Northern hybridisation, and to identify the presence of and/or to
 CC determine the degree of similarity between two (or more) nucleic acid
 CC sequences
 XX
 SQ Sequence 282 BP; 86 A; 50 C; 71 G; 75 T; 0 U; 0 Other;

Query Match 16.9%; Score 266.2; DB 6; Length 282;
 Best Local Similarity 98.6%; Pred. No. 3.7e-64;
 Matches 279; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 86 CGAGTACAAATCTTTGTGAGCGGGAAATGGCGCATGATGAGCGTCAACCTACCATC 145
 Db 1 CGAGTACAAATCTTTGTGAGCGGGAAATGGCGCATGATGAGCGTCAACCTACCATC 60
 QY 146 TGGGAGTTTGGCATAGTTAAACACACTTCTGACGAGGGAATATACCAATTAACAC 205
 Db 61 TGGGAGTTTGGCATAGTTAAACACACTTCTGACGAGGGAATATACCAATTAACAC 120
 QY 206 CTTATCAAGTCCAAAGTCACTGCAAGCAGCATGATGATGATGATGATGATGATGAT 265
 Db 121 CTTATCAAGTCCAAAGTCACTGCAAGCAGCATGATGATGATGATGATGATGATGAT 180
 QY 266 TCAACGTTACGTTACGTTGTGAGTGGCAGCGTTTCAAGAGGTACTCTGAGTTTCA 325
 Db 181 TCAACG-GGGGTTACGTTGTGAGTGGCAGCGTTTCAAGAGGTACTCTGAGTTTCA 239
 QY 326 GCGTGCATACAAATATCTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 368
 Db 240 GCGTGCATACAAATATCTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 282

RESULT 8
 ID AAA51357 standard; cDNA; 514 BP.
 AC AAA51357;
 DT 26-SEP-2000 (first entry)
 DE Wheat sucrose non-fermenting 4 protein cDNA from clone wreln.pk0143.e2.
 XX Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 KW carbon catabolite repression; development; nitrogen partitioning; ss.
 XX Triticum aestivum.
 XX Key Location/Qualifiers
 FH 19..249
 FT /*tag= a
 FT /*product= "sucrose_nor-fermenting_4_protein"
 FT /*partial
 XX WO2000036116-A2.
 XX 22-JUN-2000.

XX 15-DEC-1999; 99WO-US029825.
 PF 16-DEC-1998; 98US-0112564P.
 PR (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Allen SM, Heppard EP, Miao G, Weng Z;
 XX WPI; 2000-431594/37.
 DR 2-PSDB; AAY96793.
 XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 XX involved in carbon catabolite repression in plants and seeds, useful for
 XX controlling carbon and nitrogen partitioning pathways during plant growth
 XX and development.
 PS Claim 18; Page 47; 48pp; English.
 CC This cDNA encodes a partial wheat sucrose non-fermenting 4 (SNF4) protein
 CC which is involved in carbon catabolite repression in plants and seeds.
 CC The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae. The polynucleotides
 CC are used in plants to control carbon and nitrogen partitioning pathways
 CC during plant growth and development. The catabolite repression pathways
 CC would facilitate studies for better understanding the mechanism of
 CC catabolite repression in plants and could provide genetic tools to
 CC enhance or otherwise alter the accumulation of carbohydrates, lipids and
 CC proteins during plant growth and development
 XX
 SQ Sequence 514 BP; 141 A; 125 C; 104 G; 139 T; 0 U; 5 Other;

Query Match 16.4%; Score 259; DB 3; Length 514;
 Best Local Similarity 75.5%; Pred. No. 5.1e-62;
 Matches 358; Conservative 0; Mismatches 110; Indels 6; Gaps 3;
 QY 670 GACCGCAACAGCAGCTCTAGTGCATGCCACCCCTTATAGTCTCTGAGGACATTCGAGTAA 729
 Db 5 GATCAATCAGCATCTAGTGCATGCCACCCCTTATAGTCTCTGAGGAGTATTCGCATGA 64
 QY 730 AGCTTTTGAAATAGGCAATTTTACAGTGCAGTATTTATTTATTCATCATCATCATCAT 789
 Db 65 AAATACCTCGAATCTGGCATTTCTACAGTCCCAATCATCTATTCATTCATCATCATCAT 124
 QY 790 CATTCGCGCAGTTATTCATCTTCATCATCATCTTCGAAATTTTGAATGTATTTGTAGAT 849
 Db 125 CGTTTCGCGAGCTGTTGCACTTCGATCCCTTTTCAAGATTTTGAATGTATTCGATAGAT 184
 QY 850 ACTTCAAAAACCTCAACTGGTAAATTTGGCTATTTCTGAAACCAACCAAGTGTCTCATTCGCG 909
 Db 185 ACTTCAAGAACTCCACTGGTAGTTTGGCGATTCTTAAACCAACCAAGTGTCTCATTCGCG 244
 QY 910 TGGGT--CCTGGGTTCCGAAATTTGGTGTATCTGACAGTCTGCTCATTTGGCTGTGTC-C 966
 Db 245 TGGGGTACCTGGGGTTCGAAATTTGGTGTATCTGACAGTCTGCTCATTTGGCTGTGTC 304
 QY 967 GACCTAATGCTCACTTAGCTCTCCCTTTAACTATTTGGTTTCAAGCTGGAGTAAAGCTCAA 1026
 Db 305 GGCCTAATACATCTCTTAACTCTGCTTAACTATTTGGTTTCAAGCTGGAGTAAAGCTCAA 364
 QY 1027 TACCAAT---TGTGGATGACAAACGACTCCCTGCTTGACACTTACTCTAGAGTGCATCA 1083
 Db 365 TACCCATTGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424
 QY 1084 CAGCCCTAGCAAAAGACAGGTCTTACACATGTTTCGGCTGATGATGATGATGATGATGAT 1137
 Db 425 CANTCTGCGAAAGNAAGGCTACACCTACCGCTAGATGATGATGATGATGATGATGATGAT 478

RESULT 9
 ID AAA51349 standard; cDNA; 702 BP.
 XX

AC	AAA51349;
XX	
DT	26-SEP-2000 (first entry)
XX	
DE	Rice sucrose non-fermenting 4 protein cDNA from clone rls6.pk0037.g9.
XX	
DE	Rice; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
KW	carbon catabolite repression; development; nitrogen partitioning; ss.
XX	
OS	Oryza sativa.
XX	
PH	Key Location/Qualifiers
FT	CDS 3..512
FT	/**tag= a
FT	/label= SNF4
FT	/product= "sucrose_non-fermenting_4_protein"
FT	/partial
XX	
PK	WC200036116-A2.
XX	
PD	22-JUN-2000.
XX	
PF	15-DEC-1999; 99WO-US029825.
XX	
PR	16-DEC-1998; 98US-0112564P.
XX	
PA	(CUPO) DU PONT DE NEMOURS & CO E I.
XX	
PI	Allen SM, Heppard EP, Miao G, Weng Z;
XX	
DR	MP1; 2000-431594/37.
XX	
PT	New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
PT	involved in carbon catabolite repression in plants and seeds, useful for
PT	controlling carbon and nitrogen partitioning pathways during plant growth
PT	and development.
XX	
PS	Claim 2; Page 35-36; 48pp; English.
XX	
CC	This cDNA encodes rice (<i>Oryza sativa</i>) sucrose non-fermenting 4 (SNF4)
CC	protein which is involved in carbon catabolite repression in plants and
CC	seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
CC	<i>Arabidopsis thaliana</i> and <i>Saccharomyces cerevisiae</i> , from clone
CC	rls6.pk0037.g9 prepared from rice leaf 15 days after germination, 6 hours
CC	after infection of strain Magporthe grisea 4360-R-67 (AVR2-YAMO)
CC	(susceptible). The polynucleotides are used in plants to control carbon
CC	and nitrogen partitioning pathways during plant growth and development.
CC	The catabolite repression proteins would facilitate studies for better
CC	understanding the mechanism of catabolite repression in plants and could
CC	provide genetic tools to enhance or otherwise alter the accumulation of
CC	carbohydrates, lipids and proteins during plant growth and development
XX	
SQ	Sequence 702 BP; 233 A; 136 C; 155 G; 178 T; 0 U; 0 Other;
Query Match 9.4%; Score 148.4; DB 3; Length 702;	
Best Local Similarity 55.9%; Pred.No. 5.9e-31;	
Matches 315; Conservative 0; Mismatches 231; Indels 18; Gaps 1;	
Qy	779 ATCAGATGGATCAATCCGCGAGTATTTCGCATCTTGATCATCTTTCTGGAAATTTGAATAATG 838
Dd	14 ACCGGATTATCATCGGATGCCCTTGCTGGTCTTGCAACCCTCCAGGATTTGGAATT 73
Qy	839 TAATTTGTAGTACTTCAAANAACATCACTGGTAATTTGCTATTCGAAACCAACCGTGTG 898
Dd	74 TAATTTGCTCAAAGCTGCAAGAACAAGCCTGAAGGGTACTCATTTCTGCAAAACCAATTTG 133
Qy	899 CTCATTCCGGTGGGTTCTGGGTTCCGAAATTTGGTATCTGAACAGTCGTCATTGGC 958
Dd	134 CAGTATGSCCTATTGGTAGATGGTACCACTACTGCGAAGCGACAGATAGACAGCTTAG 193
Qy	959 TAATTTGCGAAGCTTAATGCTCACTAGCTCTGCCCTTAACATGTTGGTTCAGCTGGAGT 1018
Dd	194 AACHTGCGACCAAGCATCTCTTAAATTCATGCTGGAATTTGCTTGAAGATAGAGT 253

Qy	1019	AACTCAATACCAATTTGTGATGACAAAGCACTCGCTGCTTTCACCTTACTCTAGAACTGA	1078
Db	254	AAAGCTCAATTCCTATAGTTGACGATAATGGCGCTCTCCTTTGATGTCTACTCGCTCAGTGA	313
Qy	1079	CATCACAGCCCTTAGCAAAAGACAAAGGCTCTACACATGTTTCGGCTGCATCAGATGACCAT	1138
Db	314	TATCATGGCTCTAGGCAAGAAATGATGCTACATCTGTAATTCAGCTTGAACAGGTGACGGT	373
Qy	1139	TCACCAAGCTTTTCAGCTTCGGACAAGATGCCAATPACACCTTTTGGATTCTTTAAACGGCCA	1198
Db	374	GGAGCATGCTTCGAGCTGCAATACCAAGGTGAATGGCC-----GAAG	415
Qy	1199	GAGATGCCAGATGTGCTCGGCTGTGATCTCTTTGCTGAAGGTGATGAGCGACTGGCTAA	1258
Db	416	ACACTGTCTACCTGCTTGAGCACTAGTACCTCTCTGGAGGTTTGGAGCAATTGTGCGC	475
Qy	1259	TCCTGGGGTGGCGGGGTGTTTCATTGTCGAAGCTGGGAGCAACCTGTGAGGGGCATCAT	1318
Db	476	TCCAGGGGTGGCGGAGTGTGTTTATTTGACCAAGGAGCAGATTGTGCAAGGAATAT	533
Qy	1319	ATCACTAAGTGACATTTTCAAGTT	1342
Db	536	CTCATTGAGGGAGCGCAATTTACATT	559

RESULT 10

ABL75404

ID ABL75404 standard; cDNA; 178 BP.

XX ABL75404;

XX

DT 14-MAY-2002 (first entry)

XX

DE Corn tasse2-derived polynucleotide (cdps) SEQ ID NO:4778.

XX

KW Corn; corn tassell-derived polynucleotide; cdps; hybrid breeding; CDPS;

KW inheritance; characteristic; growth; development; disease resistance;

KW environmental adaptability; quality; yield; molecular marker;

KW multigene trait; plant breeding; corn tassell; gene; ss.

XX

OS Zea mays.

OS

XX

XX US2001051335-A1.

XX

XX

XX 13-DEC-2001.

XX

XX 16-APR-1999; 99US-00294093.

XX

XX 21-APR-1998; 98US-0082567P.

XX

XX {LALG/} LALGUDI R V.

PA {ITOL/} ITO L Y.

PA {SHER/} SHERMAN B K.

XX

XX Lalgudi RV, Ito LY, Sherman BK;

PI

XX

XX WPI; 2002-163647/21.

XX

XX Novel purified corn tassell-derived polynucleotide useful for determining

PT altered gene expression, to recover regulatory elements and to follow

PT inheritance of desirable characteristics through hybrid breeding

PT programs.

XX

XX Claim 1; SEQ ID NO 4778; 20ipp; English.

XX

XX The present sequence describes a purified corn tassell-derived

CC polynucleotide sequence (cdps) comprising a nucleic acid sequence

CC selected from those given in ABL70827 to ABL76833. The cdps sequences

CC encode corn tassell-derived polypeptides (CDPS). The cdps sequences (I)

CC can be used for determining altered gene expression, to recover

CC regulatory elements and to follow inheritance of desirable

CC characteristics through hybrid breeding programs. (I) are also useful in

the evaluation, and alteration of desired characteristics associated with growth and development, disease resistance, environmental adaptability, quality and yield, and as molecular markers for studying inheritance of multigene traits in a plant breeding program. (I) can be used to produce a tassel-specific profile of gene transcription, a transcript image, to clone regulatory elements for use in transformation vectors, to express a polypeptide, to identify, isolate or extend identical or related corn tassel nucleic acid sequences from DNA libraries, in nucleic acid hybridisation or amplification technologies, as query sequences to determine homology of known sequences, as probe for use in Southern or Northern hybridisation, and to identify the presence of and/or to determine the degree of similarity between two (or more) nucleic acid sequences

Sequence 178 BP; 55 A; 40 C; 38 G; 45 T; 0 U; 0 Other;

Query Match 8.8%; Score 139; DB 6; Length 178;
Best Local Similarity 92.8%; Pred. No. 1.4e-28;
Matches 168; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

994 TTAACATGTTGGTTCAGCTGGAGTGAAGCTCAATACCAATTGTGGATGACAACTGCC 1053
1 TTAACATGTTGGTTCAGCTGGAGTGAAGCTCAATACCAATTGTGGATGAAACGACTGCC 60

1054 TGCTTGACACTTACTCTAGAGTGCATCATCAGCCCTAGCAAAAGACAGGCTCTACAC 1113
61 TGCTTGACACTTACTCTAGAGTGCATCATCAGCCCTAG-TAAAGACAGGCTCTACAC 119

1114 ATGTTGCGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
120 ATGTTGCGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 177

1174 C 1174
178 C 178

RESULT 11
ID AAA51355 standard; cDNA; 542 BP.
XX AAA51355;
AC AAA51355;
XX 26-SEP-2000 (first entry)
DE Rice sucrose non-fermenting 4 protein cDNA from clone rls6.pk0037.g9.
XX Rice; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
KW carbon catabolite repression; development; nitrogen partitioning; ss.
XX Oryza sativa.
XX Key Location/Qualifiers
FH 182..355
CDS /*tag= a
FT /*label= SNF4
FT /*product= "sucrose non-fermenting_4_protein"
FT /*partial
FT 248..250
FT /*tag= b
FT /*note= "Encodes Xaa which is undefined"
XX
XX WO200036116-A2.
XX
PD 22-JUN-2003.
XX
PP 15-DEC-1999; 99WO-US029825.
XX
PP 16-DEC-1998; 98US-0112564P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Heppard BP, Miao G, Weng Z;

WPI; 2000-431594/37.
P-PSDB; AAY96791.
New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins involved in carbon catabolite repression in plants and seeds, useful for controlling carbon and nitrogen partitioning pathways during plant growth and development.
Claim 18; Page 46; 48pp; English.
This partial cDNA encodes rice sucrose non-fermenting 4 (SNF4) protein which is involved in carbon catabolite repression in plants and seeds. The cDNA was isolated, based on similarity to SNF4 proteins from Arabidopsis thaliana and Saccharomyces cerevisiae, from clone rls6.pk0037.g9 prepared from rice leaf 15 days after germination, 6 hours after infection of strain Magaportha grisea 4360-R-67 (AYR2-YAMO) (susceptible). The polynucleotides are used in plants to control carbon and nitrogen partitioning pathways during plant growth and development. The catabolite repression proteins would facilitate studies for better understanding the mechanism of catabolite repression in plants and could provide genetic tools to enhance or otherwise alter the accumulation of carbohydrates, lipids and proteins during plant growth and development

Sequence 542 BP; 150 A; 119 C; 122 G; 148 T; 0 U; 3 Other;

Query Match 6.0%; Score 95; DB 3; Length 542;
Best Local Similarity 55.5%; Pred. No. 4.8e-16;
Matches 182; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

779 ATCAGATGGATCATTCGCGAGTTATTGCATCTTGCATCATCTTCTGGAATTTGAAATG 838
7 ACCGATTTCATCAGGATGCTCTTGTCTGCTTGCACCCCTCCAGGATTTGCAATT 66

839 TATTGTAGATATCTCAAAACTCACTGGTAAATTTGGCTTATCTGAAACCAACAGTGTG 898
67 TATTGTCTCAAAAGCTGCAAGACAGCCCTGGAAGGTACTCATTTCTGAAAACAGATTGT 126

899 CTCATTCGCTGGTTCCTGGTTCGAAATTTGGTATCTGACAGCTGCTCCATTGGC 958
127 CAGTATGCTTATTTGTTACATGTCACCATCTGCGAGGCAAGCAATAGACCTTAG 186

959 TATGTTGGAGCTAATGCTCCTCATCTAGCTCTGCTCCCTTAACATTTGTTGAGCTGGAGT 1018
187 AACTTCGGACCAAGCACTCTCTTAATTCATGCTGAGTTTGTCTTGAAGATAGCT 246

1019 AAGCTCAATACCAATTTGGATGACACAGCTCCCTGCTTGACACTTACTCTAGAGTGA 1078
247 ANGCTCAATTCCTATAGTTGACGATATGCGCTCTCCTTGATGCTACTCTCCTCAGTGA 306

1079 CATCACAGCCCTAGCAAAAGACAAGGTC 1106
307 TATCATGCTCTAGGCAAGATGATGTC 334

RESULT 12
ABS51409
ID ABS51409 standard; cDNA; 1905 BP.
XX ABS51409;
AC ABS51409;
XX 05-NOV-2002 (first entry)
XX cDNA encoding larval viability associated protein #32.
XX Fruit fly; larval viability; insecticidal activity; maize; wheat; oat;
KW rye; sorghum; rice; barley; millet; turf; cotton; sugarcane; sugar beet;
KW oilseed rape; soybean; vegetable crop; fruit; gene; ss.
XX Drosophila melanogaster.
XX WO200257455-A2.
XX

25-JUL-2002.
18-JAN-2002; 2002WO-US001568.
18-JAN-2001; 2001US-0262351P.
(SYGN) SYNGENTA PARTICIPATIONS AG.
Stam L, Bachmann J, Broadus J, Kamdar KP;
WPI; 2002-590746/63.
P-PSDB; ABG70033.
Identifying inhibitors of activity of proteins essential for Drosophila larval viability comprises expressing in a host a protein essential for larval activity and identifying compounds that inhibit or interact with the protein.
Claim 1; Page 157-158; 169pp; English.
The invention describes a method of identifying compounds that inhibit the activity of, or that interact with a protein essential for Drosophila larval viability comprising expressing in a recombinant host a DNA molecule to produce a protein essential for larval viability. The method is useful for identifying compounds with insecticidal activity. Compounds identified are useful as insecticides in crops such as maize, wheat, oats, rye, sorghum, rice, barley, millet, turf, cotton, sugarcane, sugar beet, oilseed rape, soybeans, vegetable crops and fruits. This sequence encodes a fruit fly larval viability associated protein
XX
SQ
Sequence 1905 BP; 486 A; 533 C; 521 G; 365 T; 0 U; 0 Other;
Query March 3.2%; Score 50.6; DB 6; Length 1905;
Best Local Similarity 54.6%; Pred. No. 0.0024;
Matches 101; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 371 TTGTCATACATGCTATGATTTACTCCAGATTTCGGCAAGGTTATGTCCTAGACATTAA 430
DB 489 TTTTCAAGTGCTATGATCTGATACCCACCTCCGCCAAGTTGGTTGTTCTCGACACCCA 548
QY 431 TTACCTGTGAGCAATCATTCATATCTCCATGACAGGGGATTCCTGTAGCTCCTCT 490
DB 549 GCTTCTGTGAAGAGGCGCTTCTAGCCCTCGTCTACACGGTGTGCGAGCGGCGCGCT 608
QY 491 CTGGGACTCATTCAGAGGTCAATTTGTTGGTCTCCTTAGCCCATTTGGATTTCATCTCAT 550
DB 609 CTGGGATTCGGAGAGCAACAGTCTGTGGGCTGCTAACCATCAGGACTTATCAAGAT 668
QY 551 ATTGC 555
DB 669 CCTGC 673
RESULT 13
ABLI18857
ID ABLI18857 standard; DNA; 3261 BP.
XX
AC ABLI18857;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 8044.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2001; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR NPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
XX
PS Claim 1; SEQ ID NO 8044; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLI16178-ABLI30511), expressed DNA sequences (ABLI01840-ABLI16175), and the encoded proteins (ABBI5773-ABBI72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3261 BP; 810 A; 952 C; 931 G; 568 T; 0 U; 0 Other;
Query Match 3.2%; Score 50.6; DB 4; Length 3261;
Best Local Similarity 54.6%; Pred. No. 0.0031;
Matches 101; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 371 TTGTCATACATGCTATGATTTACTCCAGATTTCGGCAAGGTTATGTCCTAGACATTAA 430
DB 1780 TTTTCAAGTGCTATGATCTGATACCCACCTCCGCCAAGTTGGTTGTTCTCGACACCCA 183
QY 431 TTACCTGTGAGCAATCATTCATATCTCCATGACAGGGGATTCCTGTAGCTCCTCT 490
DB 1840 GCTTCTGTGAAGAGGCGCTTCTAGCCCTCGTCTACACGGTGTGCGAGCGGCGCGCT 189
QY 491 CTGGGACTCATTCAGAGGTCAATTTGTTGGTCTCCTTAGCCCATTTGGATTTCATCTCAT 490
DB 1900 CTGGGATTCGGAGAGCAACAGTCTGTGGGCTGCTAACCATCAGGACTTATCAAGAT 195
QY 551 ATTGC 555
DB 1960 CCTGC 1964
RESULT 14
ABLI18856
ID ABLI18856 standard; DNA; 39651 BP.
XX
AC ABLI18856;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 8041.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2001; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.

XX FA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 8041; 21pp + Sequence Listing; English.
XX PS The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL1840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 39651 BP; 11721 A; 8330 C; 8039 G; 11561 T; 0 U; 0 Other;

Query Match 3.2%; Score 50.6; DB 4; Length 39651;
Best Local Similarity 54.6%; Pred. No. 0.0095;
Matches 100; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 371 TTTCATACATGCTATGATTACTCCAGATTCTGGCAAGGTATTGGCCCTAGACATTAA 430
Db 36985 TTTCACAGAGCTATGATCTGTATACCACTCCGCCAAGTGTGTTCGACACCCA 36944
QY 431 TTACCTGTGAAGCAATCATTCATATTCTCCATGAACAGGGGATTCCTGTAGCTCTCT 490
Db 36945 GCTTCTGTAAAGAAGGCTTCTAGCCCTCTGTCTACAAAGGTGTGGAGCGGACCGCT 37004
QY 491 CTGGAGCTCATTCAGAGTCAATTGTTGGTCTCTCTTAGCCCATGGATTTTCATCTCAT 550
Db 37005 CTGGATTCGGAGAGCAACAGTTCGTGGGATGCTAACCATCAGCGGACTTTATCAAGAT 37064
QY 551 ATTGC 555
Db 37065 CCTGC 37069

RESULT 15
ACC60984
ID ACC60984 standard; DNA; 969 BP.
XX AC ACC60984;
XX AC
XX DT 20-JUN-2003 (first entry)
XX DE Gene sequence #SEQ ID 750.
XX KW Multiprotein complex; eukaryote; drug target; diagnosis; gene; ds.
XX OS Saccharomyces cerevisiae.
XX PN EPI258494-A1.
XX PD 20-NOV-2002.
XX PF 20-DEC-2001; 2001EP-00130253.
XX PR 15-MAY-2001; 2001EP-00111774.
XX PA (CELL-) CELLZOME AG.
XX PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
PI Marzioch M, Schultz JD, Superti-Furga GD;

XX WPI; 2003-250078/25.
XX P-PSDB; ABR52942.
XX PT New isolated protein complexes useful for diagnosing a disease or
XX PT disorder, or as a target for an active agent of a pharmaceutical,
XX PT preferably a drug target in the treatment or prevention of disease or
XX PT disorder.
XX PS Disclosure; SEQ ID NO 750; 17pp + Sequence Listing; English.
XX PS The invention relates to multiprotein complexes from eukaryotes. Proteins
CC of the invention and DNA sequences encoding them are given in records
CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
CC obtainable by using a protein as a bait and isolating the set of proteins
CC which is attached thereto from cells. Such protein complexes may comprise
CC up to 30 distinct proteins. Protein complexes of the invention are useful
CC for diagnosing a disease or disorder, or as a target for an active agent
CC of a pharmaceutical, preferably a drug target in the treatment or
CC prevention of a disease or disorder. Note: the sequence data for this
CC patent is not represented in the printed specification, but is based on
CC sequence information supplied by the European Patent Office. The complete
CC document is available on CD-ROM
XX SQ Sequence 969 BP; 299 A; 173 C; 203 G; 294 T; 0 U; 0 Other;

Query Match 3.1%; Score 49.6; DB 7; Length 969;
Best Local Similarity 54.3%; Pred. No. 0.0034;
Matches 100; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 361 AATATCTGAATTTGCATACATGCTATGATTACTCCAGATTCTGGCAGGTTATTGCC 420
Db 71 AGTTTTGAATCTGAAACATCTTATGACGTGTTCCTGTCTTACCGTTAATTTGCT 130
QY 421 TAGACATTAAATTTACCTGTGAAGCAATCATTCATATTCCTCAACAGGGGATTCCTG 480
Db 131 TGGACACCTCGTGTGTAGTGAAGAAATCACTGAATGTTCTTTGCAAAATAGCATGTCT 190
QY 481 TAGCTCTCTCTCTGGGACTCATTCAGAGGTCAATTTGTTGGTCTCTTACCCCATTTGATT 540
Db 191 CTGGCCATTATGGGACTCCAAAGACTTCCAGGTTCGCTGGACTTCTACTACTACAGATT 250
QY 541 TCAT 544
Db 251 TTAT 254

Search completed: July 9, 2004, 05:16:45
Job time : 723 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 03:22:12 : Search time 4369 Seconds
(without alignments)
10771.984 Million cell updates/sec

Title: US-09-857-525C-1

Perfect score: 1576

Sequence: 1 gcacgaggggaacatttgcg.....tcaaaaaaaaaaaaaa 1576

Scoring table: IDENTITY NJC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

BST.*

1: em_estba.*

2: em_estnum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_esti.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hur.*

18: em_gss_inv.*

19: em_gss_pin.*

20: em_gss_vit.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vri.*

28: gb_gsa1.*

29: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1200.6	76.2	2227	11	AY109512
2	689.6	43.8	873	14	CD434895
3	645.8	41.0	690	13	BQ163922
4	642.8	40.8	689	14	CF635900

C	5	629.4	39.9	663	14	CF011159
C	6	627.8	39.8	664	14	CF011157
	7	607.4	38.5	627	9	AW053113
	8	596.8	37.9	623	14	CF011158
	9	581.2	36.9	644	10	AW744961
C	10	580.4	36.8	601	13	BU098591
	11	575.6	36.5	616	14	CF627465
	12	574	36.4	666	14	CA211312
	13	572.8	36.3	683	14	CA262684
	14	560.4	35.6	574	14	CF011160
	15	557.4	35.4	559	10	BE123279
	16	552.8	35.1	631	14	CA253676
	17	547.4	34.7	549	13	BU499410
	18	539	34.2	738	14	CD878389
	19	537.4	34.1	595	10	BE598304
	20	520.8	33.0	534	12	BI423527
	21	520.6	33.0	680	13	CA148077
	22	518.4	32.9	712	14	CD903099
	23	516.4	32.8	737	14	CA500333
C	24	512.8	32.5	558	12	BI388981
	25	511.4	32.4	577	14	CA218557
	26	503.6	32.0	558	10	BE598914
C	27	498.2	31.6	566	14	CF624152
C	28	494	31.3	503	10	AW313249
	29	493.4	31.3	911	13	CA067477
C	30	492.6	31.3	511	9	AI987202
	31	491.8	31.2	596	14	CF052805
	32	474.6	30.1	700	13	CA124616
C	33	465	29.5	831	10	BE704511
	34	463.8	29.4	626	14	CD995954
	35	456.4	29.0	594	12	BJ450369
	36	448.4	28.5	726	14	CD902160
	37	441.4	28.0	443	9	AW056027
	38	441.2	28.0	604	12	BJ477917
	39	434.6	27.6	563	12	BG605114
	40	428	27.2	473	10	AW676933
	41	422.6	26.8	612	14	CA595438
	42	420.8	26.7	626	14	CA498139
	43	409.2	26.0	693	14	CD863501
	44	408	25.9	524	12	BG560098
	45	404.6	25.7	549	13	BU969888

ALIGNMENTS

RESULT 1

AY109512

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Zea mays

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 2227)

Rainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes

Unpublished (2002)

2 (bases 1 to 2227)

Coe, E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones, these are publically available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.lscate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

AY109512

Zea mays

CL367_1 mRNA sequence.

2227 bp

linear

HTC 17-OCT-2002

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB; www.zmdb.iastate.edu.

FEATURES

source

location/Qualifiers

1..2227

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="MaizeDB:631725"

/db_xref="taxon:4577"

/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

```

Query Match      76.2%; Score 1200.6; DB 11; Length 2227;
Best Local Similarity 90.8%; Pred. No. 6.3e-222;
Matches 1239; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

Qy 8  GGAACATTTGCGGATGTCCTCTATCGAAGGCTGCCCACTGTATTTCAAGGCTATTTGCAG 67
Db 639 GGAACATTTGCGGATGTCCTCTGTCGAAGGCTGCCCACTGTATTTCAAGGCTATTTGCAG 698

Qy 68  TCTGTCTCTGGGATTCACGATGACAAATCTTTGTGGACGGGAAATGCGCGCATGATGA 127
Db 699 CTTGTCTCCAGGATTCATGATGACAAATCTTTGTGGACGGGAAATGCGCGCATGATGA 758

Qy 128 GCGTCAACCTACCATATCTGGGAGTTGGCATAGTTAAACACATTTACTTTGACAAAGGA 187
Db 759 GCGTCAACCTACTATCTGGGAGTTGGCATAGTTAAACACATTTACTTTGACAAAGGA 818

Qy 188 ATATAACCAATAAACCTTTATCAAGTCCAAAGCACACCTGGAAGCAGGATGAACATGGA 247
Db 819 ATTTAATCAATAAATGCTTTATTAATCCAAAGCACACCTGGAAGCAGGATGAACATGGA 878

Qy 248 TGTGGTAATGAATAATTTCAAGTACGGTTACGTTTCAGATGCGACCGCTTTCAGAGG 307
Db 879 TGTGGTAATGAATAATTTCAAGTACGGTTACGTTTCAGATGCGACCGCTTTCAGAGG 938

Qy 308 TACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTCTTCTGAATATCT 367
Db 939 TACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTCTTCTGAATATCT 998

Qy 368 GAATTTGCATACATGCTATGATTTATCTCCAGATTCGGCAAGGTTATTCGCCCTAGACAT 427
Db 999 GAATTTGCATACATGCTATGATTTATCTCCAGATTCGGCAAGGTTATTCGCCCTAGACAT 1058

Qy 428 TAATTTACCTGTGAAGCAATCATTCATATTCCTCATGACAGGAGTTCTGTTAGTCTCC 487
Db 1059 TAATTTACCTGTGAAGCAATCATTCATATTCCTCATGACAGGAGTTCTGTTAGTCTCC 1118

Qy 488 TCTCTGGAGCTATTTCAGAGTCAATTTGTTGTTCTCTTAGCCCATTTGATTTTCACT 547
Db 1119 TCTCTGGAGCTATTTCAGAGTCAATTTGTTGTTCTCTTAGCCCATTTGATTTTCACT 1178

Qy 548 CATATTTGGGAGCTAGAACTCATGCTCGAATTCGACAGAGAGGCTTGAACACACA 607
Db 1179 TATATTTGGGAGCTAGAACTCATGCTCGAATTCGACAGAGAGGCTTGAACACACA 1238

Qy 608 CACTATATCTGCATGGAAGAGGCTTAAGCGCAAACTAAATGGAAGAATGATAGTCAGTG 667
Db 1239 CACTATATCTGCATGGAAGAGGCTTAAGCGCAAACTAAATGGAAGAATGATAGTCAGTG 1298

Qy 668 GCGACCCAAACAGCATCTAGTGCATGCCACCCCTTATGATTCCTTGGAGGACATTCAGT 727
Db 1299 GCGAGCACATCAGCATCTAGTGCATGCCACCCCTTATGATTCCTTGGAGGACATTCAGT 1358

Qy 728 AAAGCTTTTGCATAATGGCATTTCTACAGTGCAGGTTATTTATTCATCATCATCATGATGG 787

```

```

Db 1359 AAAGCTTTTGTCTAATGACATTTCTACAGTGCCAGTTATTTATTCATCATCATCATGAG 1418
Qy 788 ATCATTCCTCCGAGTTATTCATTCCTTCATCATCTTTCTCTGAAATTTTGAATGATTGTAG 847
Db 1419 ATCATTCCTCCGAGTTATTCATTCCTTCATCATCTTTCTCTGAAATTTTNNNNNNNNNN 1478
Qy 848 ATACTTCAAAACCAACTGCTGTAATTTGCTGATCTGAACCAACCAAGTGTGCTCCATTC 907
Db 1479 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1538
Qy 908 GCTGGTTCCTGGGTTCCGAAATTTGGTGAATCTGAACAGTCTCTCCATTCGCTATGCTG 967
Db 1539 GCTGGTTCCTGGGTTCCGAAATTTGGTGAATCTGAACAGTCTCTCCATTCGCTATGCTG 1538
Qy 968 ACCTAATGCTCATCTAGTCTGCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1327
Db 1599 ACCTAATGCTCATCTAGTCTGCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1658
Qy 1028 ACCAATTTGCTGATGACACAGTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1087
Db 1659 ACCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1718
Qy 1088 CCTAGCAAAAGCAAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACAGGC 1147
Db 1719 CCTAGCTAAAGCAAGGCTTACACATGTTGGCTGGATGAGATGAGATGAGATGAGATG 1778
Qy 1148 TTTGAGCTTGGACAGATGCAATACACATTTTGGATCTTTAAGCGGACAGATGCA 1207
Db 1779 TTTAAGCTTGGACAGATGCAATACACATTTTGGATCTTTAAGCGGACAGATGCA 1838
Qy 1208 GATGCTCTCGGCTCTGATCTTCTGCTGAAAGTATGAGGACATGCTGCTGCTGCTGCTG 1267
Db 1839 GATGCTCTCGGCTCTGATCTTCTGCTGAAAGTATGAGGACATGCTGCTGCTGCTGCTG 1898
Qy 1268 GCGCGGCTGTTCTATTTGGAAGCTGGAGCAAAAGCTGGAGGAGCATCATATCACTAAG 1327
Db 1899 GCGCGGCTGTTCTATTTAGAGAGCTGGAGCAAAAGCTGGAGGAGTATATATCACTGAG 1958
Qy 1328 TGACATTTCAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1371
Db 1959 TGATATTTCAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2002

```

RESULT 2

CD434895

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Sequencing of the maize endosperm ESTs

Unpublished (2002)

Contact: Lai, Jinsheng

Dr. Joachim Messing's lab

Waksman Institute, Rutgers University

190 Frelinghuysen Rd., Piscataway, NJ 08854, USA

Tel: 732-445-3801

Fax: 732-445-5735

Email: jlai@waksman.rutgers.edu

Seq primer: T3.

FEATURES

source

1..873

/organism="Zea mays"

/mol_type="mRNA"

/cultivar="W22"

```
ORIGIN
Query Match 43.8%; Score 689.6; DB 14; Length 873;
Best Local Similarity 92.5%; Pred. No. 3.9e-123;
Matches 736; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

/ab_xref="taxon:4577"
/tissue_type="Endospore of 7-23DAP"
/clone_lib="Endospore 3"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

6 AGGGAACATTTGCCGATGCTCTCCATCGAAGGCTGCCACCTGATTTTTCAGGCTATTGTC 65
78 AAGGAACATTTGCCGATGCTCTCCATCGAAGGCTGCCACCTGATTTTTCAGGCTATTGTC 137
66 AGCTGTCTC-CTGGGATTCAGAGTACAAATCTTTGTGACGGGGAATGGCGCATGA 124
138 AGCTGTCTC-CTCAAGATTCAGAGTACAAATCTTTGTGACGGGGAATGGCGCATGA 197
125 TGAGCGTCAACCTACCATATCTGGGAGTTTGGCATAGTTTAAACACACTTACTTTGACAAG 184
198 TGAGCGCAACCTACTATATCTGGGAGTTTGGCATAGTTTAAACACACTTACTTTGACAAG 257
185 GGAATATAACCAATTAACACCTTATCAAGTCCAGACACCTGGAAGCAGATGAACAT 244
258 GGAATTTAATCAATTAACACCTTATTAATTAACACACCTGGAAGCAGATGAACAT 317
245 GGATGTGATTAATGAAATTTTCAACGTACGGTTTACGTTGTGATGTCAGATGCAACGTTTCAGA 304
318 GGATGTGATTAATGAAATTTTCAACGTACGGTTTACGTTGTGATGTCAGATGCAACGTTTCAGA 377
305 AGTACTCTGAGAGTTTTCAGAGGCTGCAATTAACATTAATATCTAGGTTGCTGTTCTGAATA 364
378 AGTACTCTGAGAGTTTTCAGAGGCTGCAATTAACATTAATATCTAGGTTGCTGTTCTGAATA 437
365 TCTGAATTTTCATACATGCTATGATTTTACCTCCAGATTTCTGCAAGGATTTATGCGCTAGA 424
438 TCTGAATTTTCATACATGCTATGATTTTACCTCCAGATTTCTGCAAGGATTTATGCGCTAGA 497
425 CATTAATTTTCTGTGAGGAGTCAATTCATATTTCTCCATGAAAGGAGTTCCTGTAGC 484
498 CATTAATTTTCTGTGAGGAGTCAATTCATATTTCTCCATGAAAGGAGTTCCTGTAGC 557
485 TCCTCTCTGAGGAGTCAATTCATATTTCTCCATGAAAGGAGTTCCTGTAGC 544
558 TCCTCTCTGAGGAGTCAATTCATATTTCTCCATGAAAGGAGTTCCTGTAGC 617
545 ACTCATATTTGCGGAGTCAATTCATATTTCTCCATGAAAGGAGTTCCTGTAGC 604
618 ACTCATATTTGCGGAGTCAATTCATATTTCTCCATGAAAGGAGTTCCTGTAGC 677
605 ACACATATTTGCGGAGTCAATTCATATTTCTCCATGAAAGGAGTTCCTGTAGC 664
678 ACACATATTTGCGGAGTCAATTCATATTTCTCCATGAAAGGAGTTCCTGTAGC 737
665 GTGGGACCGACACATCTAGTATGATGCAACCCCTTATGATCTCTTGGGAGCATTCG 724
738 GTGGGACCGACACATCTAGTATGATGCAACCCCTTATGATCTCTTGGGAGCATTCG 797
725 AGTAAAGCTTTTGCATATGCGGATTTCTACAGTGGCAGTTTATTTATTCATCATCAG 784
798 AGTAAAGCTTTTGCATATGCGGATTTCTACAGTGGCAGTTTATTTATTCATCATCAG 857
785 TGGATCATTTCCCGCAG 800
858 TGGATCATTTCCCTCAG 873
```

RESULT 3
BQ163922/c
LOCUS
DEFINITION
952081C10.x1 952 - BMS tissue from Walbot Lab (reduced rRNA) Zea
mays cDNA, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BQ163922
BQ163922.1 GI:20300979
EST.
Zea mays
Zea mays

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Maize ESTs from various cDNA libraries sequenced at Stanford University
Unpublished (1995)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 952081 row: C column: 10.
Location/Qualifiers
1. .690
/organism="Zea mays"
/mol_type="rRNA"
/cultivar="BMS (Black Mexican Sweet)"
/db_xref="taxon:4577"
/tissue_type="suspension culture"
/dev_stages="mixed logarithmic and stationary growth phases"
/lab_host="DH10B"
/clone_lib="952 - BMS tissue from Walbot Lab (reduced rRNA)"
/notes="Vector: pUC19; Site_1: EcoRI; Site_2: EcoRI; The library was prepared by George Rudenko using poly (A) selected RNA and Universal Riboclone cDNA Synthesis System (Promega). cDNA was synthesized using both random and oligo(dT) primers in separate reactions and equipped with EcoRI adaptors. Library was size-fractionated on agarose gels (for insert size >400bp) and non-directionally cloned into EcoRI-digested pUC19 vector. Blue/white selection on carbenicillin-containing plates was used to recover positive clones."

FEATURES
source

Query Match 41.0%; Score 645.8; DB 13; Length 690;
Best Local Similarity 98.1%; Pred. No. 1.3e-114;
Matches 688; Conservative 0; Mismatches 2; Indels 11; Gaps 3;

834 AAATGTATTTGTAGATCTTCAAAAACCTCAACTGTAATTTGCTATTCTGAACCAACCA 893
690 AAATGTATTTGTAGATCTTCAAAAACCTCAACTGTAATTTGCTATTCTGAACCAACCA 631
894 GTGTGCTCCATTTCCGCTGGGTTCTCGGTTCCGAAATTTGGTGTATCTGAACAGTGTCCA 953
630 GTGTGCTCCATTTCCGCTGGGTTCTCGGTTCCGAAATTTGGTGTATCTGAACAGTGTCCA 571
954 TTGGCTATGTTGGAGCTTATGCTCACTTAGCTCTGCTGCTTCAACATGTTGGTTCAGCT 1013
570 TTGGCTATGTTGGAGCTTATGCTCACTTAGCTCTGCTGCTTCAACATGTTGGTTCAGCT 511
1014 GGAGTAAAGCTCAATACCAATTTGGGATGCAACAGCTCCCTGCTTGACACTTACTCTAGA 1073
510 GGAGTAAAGCTCAATACCAATTTGGGATGCAACAGCTCCCTGCTTGACACTTACTCTAGA 451
1074 AGTGACATCAGCCCTAGCAAAAGCAAGTGTACACATGTTCCGGTGGATGATG 1133
450 AGTGACATCAGCCCTAGCAAAAGCAAGTGTACACATGTTCCGGTGGATGATG 391
1134 ACCATTACAGGCTTTGCGAGCTTGGACAGATGCCAATACACTTTGGATTTCTTAAC 1193
390 ACCATTACAGGCTTTGCGAGCTTGGACAGATGCCAATACACTTTGGATTTCTTAAC 331

```

QY 1194 GGCAGAGATCCAGATGCTGCTCGGTCATGCTTTCGTGAGTGTGAGCGCATG 1253
Db 330 GGCAGAGATCCAGATGCTGCTCGGTCATGCTTTCGTGAGTGTGAGCGCATG 271
QY 1254 GCTAATCTCTGGGTGCGCGGGTGTTCATTGTGAAGCTGGGAGCAAACTGTGTGAGGCG 1313
Db 270 GCTAATCTCTGGGTGCGCGGGTGTTCATTGTGAAGCTGGGAGCAAACTGTGTGAGGCG 211
QY 1314 ATCATATCATTAGTGCATTTTCAGTTCTTCTGAGCTTGTGAGGAGAGAGGAGCATG 1373
Db 210 ATCATATCATTAGTGCATTTTCAGTTCTTCTGAGCTTGTGAGGAGAGAGGAGCATG 151
QY 1374 GCGAGCTTTTTCATACAGTAGGACCTCTGGGCTTTGGCTCTCGGATATTCGC 1433
Db 150 GCGAG-TTTTTTCATACAGTAGGACCTCTGGGCTTTGGCTCTCGGATATTCGC 92
QY 1434 TATGCTGAGGACTCAGGTTTCATGCTGTGAAGTTCCTTTTTCGTTTCGTTTTT 1493
Db 91 TATCG-----TCAGGTTTCATGCTGTGAAGTTCCTTTTTCGTTTCGTTTTT 42
QY 1494 ACCATTGGCCCTCCGTTGTACAAAACATACATCAGGT 1534
Db 41 ACCATTGGCCCTCCGTTGTACAAAACATACATCAGGT 1

```

RESULT 4

```

LOCUS CF635900/CF635900
DEFINITION zmrw00.0B10-001-h03.s3 zmrw00 Zea mays cDNA, mRNA sequence.
ACCESSION CF635900
VERSION CF635900.1 GI:37397210
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays

```

```

REFERENCE 1 (bases 1 to 689)
AUTHORS Bohnert, H.; Sharp, R.E.; Springer, G.K.; Poroyko, V.; Fredrickson, M.;
Topinka, C.; Davis, G.E.; Schachtman, D.; Wu, Y.; and Nguyen, H.T.
TITLE NSF Grant DBI-0211842: Functional Genomics of Root Growth and Root
JOURNAL Signaling Under Drought
COMMENT Unpublished (2003)
Contact: Hans Bohnert
University of Illinois, Urbana-Champaign
1201 West Gregory Drive, Urbana, IL 61801, USA
Tel: 217-265-5475
Fax: 217-333-5574
Email: bohnert@life.uiuc.edu
POLYA=Yes.
Location/Qualifiers
1..689
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone_lib="zmrw00"
/note="Samples were collected in Robert E. Sharp's lab
(University of Missouri-Columbia) to construct three
normalized cDNA libraries. Dark-grown maize seedlings with
primary roots 12-20 mm in length were transplanted to high
(-0.03 MPa) or low water potential (-1.6 MPa) vermiculite,
and harvested at 5 h and 48 h after transplanting. About
1,000 roots were used for each of the low water potential
libraries (zmrw05 and zmrw48) while 500 roots were
combined from each of the two time points at high water
potential (zmrw00). Each root was divided into 4 segments
(distances are from the junction of the root apex and root
cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7
mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details
of conditions see (1) with nutrient modifications as in
(2)). The three normalized cDNA libraries were constructed
in the lab of Hans Bohnert (University of Illinois-UC).

```

FEATURES

```

source
Query Match 40.8%; Score 642.8; DB 14; Length 689;
Best Local Similarity 96.4%; Pred. No. 4.8e-114;
Matches 671; Conservative 0; Mismatches 17; Indels 8; Gaps 1;
QY 850 ACTCAAAACCTCAACCTGTAATTTGCCATTTCGAACCAACCACTGCTCCATTCGCG 909
Db 689 ACTCAAAACCTCAACCTGTAATTTGCCATTTCGAACCAACCACTGCTCCATTCGCG 630
QY 910 TGGGTTCTCGGGTTCGGAATAATTGGTGATCTGAACAGTCGTCCATTCGTTGCGAC 969
Db 629 TGGGTTCTCGGGTTCGGAATAATTGGTGATCTGAACAGTCGTCCATTCGTTGCGAC 570
QY 970 CTAATGCTCACTAGCTCTGCCCTTACATCTGTTGTTCAAGCTGAGTCAAGCTCAATAC 1029
Db 569 CTAATGCTCACTAGCTCTGCCCTTACATCTGTTGTTCAAGCTGAGTCAAGCTCAATAC 510
QY 1030 CAATTGTTGATGACAAACGACTCCCTGCTTGACACTTACTTAGAAGTGACATCACAGCCC 1089
TAG_SEQ=TCGCA"
ORIGIN

```

Total RNA was extracted by the 'hot phenol' method (Plant Molecular Biology manual, DS: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (Root segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adapted with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zmrw05: 3.37 x 10⁶; zmrw48: 4.87 x 10⁶; zmrw00: 3 x 10⁶. The background of empty clones was less than 1%. Inserts ranged from ~0.5kb to >2.5 kb, as determined by PCR. Plasmid DNA from the primary libraries then was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites that flank the cloned cDNA inserts. The purified PCR products, representing the entire cDNA population cloned in each library, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30°C. Non-hybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B. The total number of clones with insert was: zmrw05: 2.0x10⁷; zmrw48: 4.2x10⁷; zmrw00: 1.1x10⁷. The background of empty clones was less than 2%. Insert size, determined by PCR of the entire library, ranged from 0.5kb to 2.5kb. (1) Sharp R.E; Silk W.K; Haio T.C. Growth of the Maize Primary Root at Low Water Potentials I. Spatial Distribution of Expansive Growth. Plant Physiology (Rockville). 87(1). 1988. 50-57. (2) Spollen W.G; LeNoble M.E; Samuels T.D; Bernstein N; Sharp R.E. Abscissic acid accumulation maintains maize primary root elongation at low water potentials by restricting ethylene production. Plant Physiology (Rockville). 122(3). March, 2000. 967-976.


```

/organism="Zea mays"
/mol_type="mRNA"
/cultivar="F2"
/db_xref="taxon:4577"
/clone="QB16f01"
/tissue_type="pollen"
/clone_lib="QB0"

ORIGIN
Query Match      39.8%; Score 627.8; DB 14; Length 664;
Best Local Similarity 99.4%; Pred. No. 3.8e-111;
Matches 651; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 905 TCCGCTGGTTCCTGGGTTCCGAAATTTGGTGAATCTGAACAGTGTCTGCAATTCGGTATGTT 964
Db 664 TCCGCTGGTTCCTGGGTTCCGAAATTTGGTGAATCTGAACAGTGTCTGCAATTCGGTATGTT 605
Qy 965 GCGACCTAATGCTCACTTACCTCTGCGCTTAAACATGTTGGTTCGAAGCTGGAGTAAGCTC 1024
Db 604 CGGACCTAATGCTCACTTACCTCTGCGCTTAAACATGTTGGTTCGAAGCTGGAGTAAGCTC 545
Qy 1025 AATACCAATTTGGATGACAAAGACTCCCTGCTTGACACTTACTCTAGAACTGACATCAC 1084
Db 544 AATACCAATTTGGATGACAAAGACTCCCTGCTTGACACTTACTCTAGAACTGACATCAC 485
Qy 1085 AGCCTAGCAAAAGACAAAGTCTACACATCTTTCGCTGGATGAGATGACCATTCACCA 1144
Db 484 AGCCTAGCAAAAGACAAAGTCTACACATCTTTCGCTGGATGAGATGACCATTCACCA 425
Qy 1145 GGCTTTTCAGCTTGCAAGATGCAATACACCTTTTGGATTCTTTAAAGCTCCAGAGATG 1204
Db 424 GGCTTTTCAGCTTGCAAGATGCAATACACCTTTTGGATTCTTTAAAGCTCCAGAGATG 365
Qy 1205 CCAGATGCTCCTCGGCTCTGATCTTCTGCTGAAGGTGATGAGGAGCTATGCTAACTCTGG 1264
Db 364 CCAGATGCTCCTCGGCTCTGATCTTCTGCTGAAGGTGATGAGGAGCTATGCTAACTCTGG 305
Qy 1265 GGTGCGCGGGTGTCTATCTGGAAGCTGGAGCAAAAGCTGTGGAGGAGCATCATCACT 1324
Db 304 GGTGCGCGGGTGTCTATCTGGAAGCTGGAGCAAAAGCTGTGGAGGAGCATCATCACT 245
Qy 1325 AAGTGACATTTCAAGTTCCTGCTGAGCTTGTGAGGAGAGAGAGCATGGCGAGTCTTT 1384
Db 244 AAGTGACATTTCAAGTTCCTGCTGAGCTTGTGAGGAGAGAGAGCATGGCGAG-TTTT 186
Qy 1385 TTAAAGTAGCAACCTCTGCGGCTTTGGGCTCTGCGGATTAATTTGGCTATGCTCAGGG 1444
Db 185 TTAAAGTAGCAACCTCTGCGGCTTTGGGCTCTGCGGATTAATTTGGCTATGCTCAGGG 126
Qy 1445 ACTCAGGTTTCATGCTCTGTAAGTTC-TTTTTTTTTGTCTGTTTTTACCAATTTGGC 1503
Db 125 ACTCAGGTTTCATGCTCTGTAAGTTC-TTTTTTTTTGTCTGTTTTTACCAATTTGGC 66
Qy 1504 CCTCCCTGTTTAAACAAACATACATCAGGTGATTTTATCTATTTGGATTTTC 1558
Db 65 CCTCCCTGTTTAAACAAACATACATCAGGTGATTTTATCTATTTGGATTTTC 11

RESULT 7
AW053113
LOCUS 614032D06.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW053113
VERSION AW053113.1 GI:5915472
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD
c-ade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 627)
AUTHORS Walbot,V.
```

```

RESULT 8
CF011158
LOCUS CF011158 623 bp mRNA linear EST 17-JUL-2003
DEFINITION QBj6f01.xg QBj Zea mays cDNA clone QBj6f01, mRNA sequence.
ACCESSION CF011158
VERSION CF011158.1 GI:32906345
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 623)
Genoplante, a major partnership french program in p-ant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com)
and http://genoplante-info.infobiogen.fr.
FEATURES
source
1..623
location/Qualifiers
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="P2"
/db_xref="taxon:4577"
/clone="QBj6f01"
/tissue_type="pollen"
/clone_lib="QBj"
ORIGIN
Query Match 37.9%; Score 596.8; DB 14; Length 623;
Best Local Similarity 99.4%; Pred. No. 3.8e-105;
Matches 620; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 905 TCGGTGGGTTCTGGGTTCCGAAATTTGGTATCTCAACAGTCGTCATGGCTATGTT 964
Db 1 TCGGTGGGTTCTGGGTTCCGAAATTTGGTATCTCAACAGTCGTCATGGCTATGTT 60
QY 965 GCGACCTAATGCTCATTAGCTCTGCCCTTAACATGTTGTTCAAGCTGGAGTAAGCTC 1024
Db 61 GCGACCTAATGCTCATTAGCTCTGCCCTTAACATGTTGTTCAAGCTGGAGTAAGCTC 120
QY 1025 AATACCAATTTGGATGACACGACCTCTGCTTGACATCTACTCTAGAAGTGACATCAC 1084
Db 121 AATACCAATTTGGATGACACGACCTCTGCTTGACATCTACTCTAGAAGTGACATCAC 180
QY 1085 AGCCCTAGCAAAAGCAAGGCTTCACACATGTTGGCTGGATGAGATGACCATCACCA 1144
Db 181 AGCCCTAGCAAAAGCAAGGCTTCACACATGTTGGCTGGATGAGATGACCATCACCA 240
QY 1145 GCTTTGACGCTTGACACAGATGCCAATACACTTTTGGATTTCTTAACGGCCAGAGATG 1204
Db 241 GCTTTGACGCTTGACACAGATGCCAATACACTTTTGGATTTCTTAACGGCCAGAGATG 300
QY 1205 CCAGATGCGCTCCGGTCTGATCTTTGCTGAAGTGATGGACGCTGGCTAACTCGG 1264
Db 301 CCAGATGCGCTCCGGTCTGATCTTTGCTGAAGTGATGGACGCTGGCTAACTCGG 360
QY 1265 GGTGCGCGGGTGTTCATTGTGGAAGCTGGGAGCAACCTGTGGGGCATCATATCACT 1324
Db 361 GGTGCGCGGGTGTTCATTGTGGAAGCTGGGAGCAACCTGTGGGGCATCATATCACT 420
QY 1325 AAGTGACATTTTCAAGTCTTCTGAGCTTGTGAGGAGAGAGCATGCGGAGTTTTT 1384
Db 421 AAGTGACATTTTCAAGTCTTCTGAGCTTGTGAGGAGAGAGCATGCGGAGTTTTT 479

```

```

QY 1385 TTAAACAGTAGCAACCCCTCTGGGCTTTGGGCTCTCGGATAATTGGCTATCGTCAGGG 1444
Db 480 TTAAACAGTAGCAACCCCTCTGGGCTTTGGGCTCTCGGATAATTGGCTATCGTCAGGG 539
QY 1445 ACTCAGGTTTCATGGTCTGTAAGTTCCTCTTTTGGTTCGTTTTCACCAATTTGGC 1503
Db 540 ACTCAGGTTTCATGGTCTGTAAGTTCCTCTTTTGGTTCGTTTTCACCAATTTGGC 599
QY 1504 CTCTCCGCTGCTTAACAAACATAC 1527
Db 600 CTCTCCGCTGCTTAACAAACATAC 623
RESULT 9
AW744961
LOCUS AW744961 644 bp mRNA linear EST 19-JUL-2000
DEFINITION LG1_385_D10.b1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION AW744961
VERSION AW744961.1 GI:7658699
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 644)
Cordonnier-Pratt M.-M., Gingle A., Marsala C. and Pratt L.H.
An EST database from Sorghum: light-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEM REV
High quality sequence stop: 621
POLYA=No.
FEATURES
source
1..644
location/Qualifiers
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/note="Organ: 10- to 14-day-old light-grown (greenhouse)
seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2:
EcoRI. The library was made from poly-A RNA in the cloning
vector lambda ZAP II. Clones to be sequenced were
prepared by mass excision."
ORIGIN
Query Match 36.9%; Score 581.2; DB 10; Length 644;
Best Local Similarity 94.1%; Pred. No. 3.8e-102;
Matches 604; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 8 GGAACATTTGCCGATGCTCTCTATCGAAGCTGCCACCTGTTTTCAGCTATTTCAG 67
Db 3 GGAACATTTGCCGATGCTCTCTATCGAAGCTGCCACCTGTTTTCAGCTATTTCAG 62
QY 68 TCTGTCTCTGGGATTCAGAGTCAAAATTCCTTGTGGACGGGATGGCGCATATGA 127
Db 63 CTGTCTCTGGGATTCAGAGTCAAAATTCCTTGTGGATGGGAGTGGCGCATATGA 122
QY 128 GCGTCAACCTACCATATCTGGGAGTTTGGCATAGTTTAAACACTTTTACTTGACAAGGA 187
Db 123 GCGCAACCTACTATCTGGGAGTTTGGATAGTTTAAACACTTATCTTGTGCAAGGA 182
QY 188 ATATACCAATAAACAACCTTATCAAGTCCAAAGCACCTCGAAGCAGGATGAACATGA 247

```



```

183 ATTTACCACTAAAGCCCTTATTAGTCCAGACACCTGGAAGCAGGATGAACATGGA 242
184
248 TGTGATAAATAAATTTTCAACGTAAGTTACGTTTACGTTTACGATGCAACCGTTTCAGAGG 307
249
243 TGTGATAAATAAATTTTCAACGTAAGTTACGTTTACGTTTACGATGCAACCGTTTCAGAGG 302
250
308 TACTCTGAGAGTTTCAGAGGTCGAATAACAATATCTAGGTTGCGTGTTCGTAATATCT 367
309
303 TACTCCGAGAGTTTCAGAGGTCGAATAACAATATCTAGGTTGCGGCGTTTCGTAATATCT 362
310
368 GAATTTGATACATCTGATGATTTTACTCCAGATTTCTGCAAGGTTTATTTGCGCTAGACAT 427
369
363 GAATTTGATACATCTGATGATTTTACTCCAGATTTCTGCAAGGTTTATTTGCGCTAGACAT 422
370
428 TAATTTACCTGTGAAGCAATCATCTCCATATTTCTCCATGAACAGGAGATTCCTGTAGCTCC 487
429
423 TAATTTACCTGTGAAGCAATCATCTCCATATTTCTCCATGAACAGGAGATTCCTGTAGCTCC 482
430
488 TCTCTGGGACTCATCTGAGAGTCATTTGTTGGTCTCTGAGCCCATTTGGATTTTCACTACT 547
489
483 TCTCTGGGACTCATCTGAGAGTCATTTGTTGGTCTCTGAGCCCATTTGGATTTTCACTACT 542
490
548 CATATTTCTGATGCAAGCAATCATCTCCATATTTCTCCATGAACAGGAGATTCCTGTAGCTCC 607
549
543 TATATTGGGAGGCTAGAACTCATGCTCGAATTTGAGAGGTCGAGAGGTCGAGAGGTCGAGAGG 602
550
608 CATATTTCTGATGCAAGCAATCATCTCCATATTTCTCCATGAACAGGAGATTCCTGTAGCTCC 649
551
603 TACGATATCTGATGCAAGCAATCATCTCCATATTTCTCCATGAACAGGAGATTCCTGTAGCTCC 644
552

```

```

RESULT 10
BU098591
LOCUS 946136F11.y1 946 - tassell primordium prepared by Schmidt lab Zea
DEFINITION 601 bp mRNA linear EST 29-AUG-2002
mays cDNA, mRNA sequence.

```

```

ACCESSION BU098591
VERSION BU098591.1 GI:22546280
KEYWORDS EST.
SOURCE Zea mays

```

```

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 601)
Walbot,V.

```

```

Maize ESTs from various cDNA libraries sequenced at Stanford
University

```

```

JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946136 row: F column: 11.
Location/Qualifiers
1. .601
/organism="Zea mays"
/mol type="mRNA"
/cultivar="OH43"
/db xref="taxon:4577"
/tissue type="tassels"
/dev stage="just after the transition from vegetative to
inflorescence development"
/lab host="XL0LR"
/clone_lib="946 - tassell primordium prepared by Schmidt
lab"
/note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;
Site 2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA

```

```

FEATURES
source

```

Library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

ORIGIN

```

Query Match 36.8%; Score 580.4; DB 13; Length 601;
Best Local Similarity 99.8%; Pred. No. 5.6e-102;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 8 GGAACATTTGGCGATGTCCTCTATCGAAGGCTGCCCACTGATTTTTCAGGCTATTTGCGAG 67
DB 20 GGAACATTTGGCGATGTCCTCTATCGAAGGCTGCCCACTGATTTTTCAGGCTATTTGCGAG 79
QY 68 TCTGTCTCTCGGATTCACGAGTACAAATTTCTTGTGACGGGGAATGGCGCATGATGA 127
DB 80 TCTGTCTCTCGGATTCACGAGTACAAATTTCTTGTGACGGGGAATGGCGCATGATGA 139
QY 128 GCGTCAACCTTACCATATCTGGGGAGTTTGGCATAGTTAAACACATTTTACTTGACAGGGA 187
DB 140 GCGTCAACCTTACCATATCTGGGGAGTTTGGCATAGTTAAACACATTTTACTTGACAGGGA 199
QY 188 ATATAACCAATTAACACCTTATCAAGTCCAGACACCTGGAAGCAGGATGAACATGGA 247
DB 200 ATATAACCAATTAACACCTTATCAAGTCCAGACACCTGGAAGCAGGATGAACATGGA 259
QY 248 TGTGATTAATCAAAATTTTCAACGTTACGTTTACGTTTACGATGCGACCGTTTCAGAGG 307
DB 260 TGTGATTAATCAAAATTTTCAACGTTACGTTTACGTTTACGATGCGACCGTTTCAGAGG 319
QY 308 TACTCTGAGAGTTTCAGAGGTCGAATAACAATATCTAGGTTGCTGTTTTCGAAATATCT 367
DB 320 TACTCTGAGAGTTTCAGAGGTCGAATAACAATATCTAGGTTGCTGTTTTCGAAATATCT 379
QY 368 GAATTTGATACATCTGATGATTTTACTCCAGATTTCTGCAAGGTTTATTTGCGCTAGACAT 427
DB 380 GAATTTGATACATCTGATGATTTTACTCCAGATTTCTGCAAGGTTTATTTGCGCTAGACAT 439
QY 428 TAAATTTACTCTGTGAAGCAATCATCTCCATATTTCTCCATGAACAGGAGATTCCTGTAGCTCC 487
DB 440 TAAATTTACTCTGTGAAGCAATCATCTCCATATTTCTCCATGAACAGGAGATTCCTGTAGCTCC 495
QY 488 TCTCTGGGACTCATCTGAGAGTCATTTGTTGGTCTCTGAGCCCATTTGGATTTTCACTACT 547
DB 500 TCTCTGGGACTCATCTGAGAGTCATTTGTTGGTCTCTGAGCCCATTTGGATTTTCACTACT 559
QY 548 CATATTTGGGAGGCTAGAACTCATGCTCGAATTTGACAGA 589
DB 560 CATATTTGGGAGGCTAGAACTCATGCTCGAATTTGACAGA 601

```

RESULT 11

```

CF627465 616 bp mRNA linear EST 02-OCT-2003
LOCUS zmrws05 OB20-011-f01.s4 zmrws05 Zea mays cDNA, mRNA sequence.
DEFINITION CF627465
ACCESSION CF627465.1 GI:37381198
VERSION EST.
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 616)
Bohnert,H., Sharp,R.E., Springer,G.K., Poroyko,V., Fredrickson,M.,
Sharp,L.G., Spollen,W.G., Riee,J., Guillen,A., Khambati,A.,
Topinka,C., Davis,G.E., Schachtman,D., Wu,Y. and Nguyen,H.T.
NSF Grant DBI-0211842; Functional Genomics of Root Growth and Root
Signaling Under Drought
Unpublished (2003)
Contact: Hans Bohnert
University of Illinois, Urbana-Champaign
1201 West Gregory Drive, Urbana, IL 61801, USA
Tel: 217-265-5475
Fax: 217-333-5574

```

Email: bohrner@life.uiuc.edu
POLYA=Yes.

FEATURES

source
1. .616
Location/Qualifiers
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone_lib="zrrws05"
/notes="Samples were collected in Robert E. Sharp's lab (University of Missouri-Columbia) to construct three normalized cDNA libraries. Dark-grown maize seedlings with primary roots 12-20 mm in length were transplanted to high (-0.03 MPa) or low water potential (-1.6 MPa) vermiculite, and harvested at 5 h and 48 h after transplanting. About 1,000 roots were used for each of the low water potential libraries (zrrws05 and zrrws48) while 500 roots were combined from each of the two time points at high water potential (zrrws00). Each root was divided into 4 segments (distances are from the junction of the root apex and root cap): segment 1, 0-3 mm plus the root cap; segment 2, 3-7 mm; segment 3, 7-12 mm; segment 4, 12-20 mm. (For details of conditions see (1) with nutrient modifications as in (2)). The three normalized cDNA libraries were constructed in the lab of Hans Bohnert (University of Illinois-UC). Total RNA was extracted by the 'hot phenol' method (Plant Molecular Biology manual, D5: 1-13, 2nd ed., 1997). This method worked in eliminating carbohydrate material present in the root tips. The integrity of the RNA was verified by denaturing agarose gels and spectrophotometry (ratio A260/280). Poly(A)+mRNA was isolated twice from total RNA using the Oligotex Direct mRNA kit (Qiagen). Poly(A)+ mRNA was converted to double-stranded cDNA and tagged by using modified Oligo(dT) primers. One of 4 sequence tags corresponding to a different segment of the root was added to the 3'-end of the modified Oligo(dT) primers, including a NotI site and used to reverse transcribe the segment-specific mRNAs into cDNAs. Each library contains all four tags. A suffix (s1, s2, s3, or s4) has been added to each sequence identifier to designate which region of the root (root segment 1, 2, 3, or 4) the sequence was found in based on the identification of the tag. A suffix of s0 indicates that the sequence tag, and hence the source segment, could not be identified. The double stranded cDNAs were size-selected (>450 bp). Size selected cDNAs were adaptor with EcoRI adaptors at both ends, and then digested with NotI. The cDNA was directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene) and electroporated into E.coli DH10B. The total number of white colony forming units (cfu) in the primary libraries before amplification was as follows: zrrws05: 3.37 x 10⁶; zrrws48: 4.87 x 10⁶; zrrws00: 3 x 10⁶. The background of empty clones was less than 1%. Inserts ranged from ~0.5 kb to >2.5 kb, as determined by PCR. Plasmid DNA from the primary libraries then was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites that flank the cloned cDNA inserts. The purified PCR products, representing the entire cDNA population cloned in each library, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 300C. Non-hybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B. The total number of clones with insert was: zrrws05: 2.0x10⁷; zrrws48: 4.2x10⁷; zrrws00: 1.1x10⁷. The background of empty clones was less than 2%. Insert size, determined by PCR of the entire library, ranged from 0.5kb to 2.5kb. (1) Sharp R E; Silk W K; Hsiao T C. Growth of the Maize Primary Root at Low Water Potential. I. Spatial Distribution of Expansive Growth. Plant Physiology (Rockville). 87(1). 1986. 50-57. (2) Spollen W G; LeNoble M E; Samuels T D; Bernstein N; Sharp R E. Abscissic acid accumulation maintains maize primary

root elongation at low water potentials by restricting ethylene production. Plant Physiology (Rockville). 122(3). March. 2000. 967-976.
TAG TISSUE=Root_segment_4
TAG_SEQ=TCGGA"

ORIGIN

Query Match 36.5%; Score 575.6; DB 14; Length 616;
Best Local Similarity 96.5%; Pred. No. 4.6e-101;
Matches 602; Conservative 0; Mismatches 14; Indels 8; Gaps 1;
Qy 935 TGATCTGAACAGTCGTCAATGGCTATGTGGACCTAAATGCTCAGTTCAGTTCGCT 994
Db 616 TGATCTGAACAGTCGTCAATGGCTATGTGGACCTAAATGCTCAGTTCAGTTCGCT 557
Qy 995 TACATCTGTGGTCAAGCTGGAGTAAGCTCAATACCAATCTGGATGCAACGACTCCCT 1054
Db 556 TACATCTGTGGTCAAGCTGGAGTAAGCTCAATACCAATCTGGATGCAACGACTCCCT 497
Qy 1055 GCTTGACACTTACTCTAGAGTGCATCAGCCCTAGCAAAAGACAAAGCTCTACACACA 1114
Db 496 GCTTGACACTTACTCTAGAGTGCATCAGCCCTAGCAAAAGACAAAGCTCTACACACA 437
Qy 1115 TGTTCGGCTGATGAGATGACCATTCACAGGCTTTCAGCTTGGACAAAGATGCCATATC 1174
Db 436 TGTTCGGCTGATGAGATGACCATTCACAGGCTTTCAGCTTGGACAAAGATGCCATATC 377
Qy 1175 ACCTTTTGGATCTTTAAACGCCAGAGATGCCAGATGCGCTCCGGCTGATCCTTTGCT 1234
Db 376 ACCTTTTGGATCTTTAAACGCCAGAGATGCCAGATGCGCTCCGGCTGATCCTTTGCT 317
Qy 1235 GAAGTGTATGAGGAGCTGGCTAATCTCTGGGGTGGGGGTGTTTCATTGTGAGCTGG 1294
Db 316 GAAGTGTATGAGGAGCTGGCTAATCTCTGGGGTGGGGGTGTTTCATTGTGAGCTGG 257
Qy 1295 GAGCAACGTTGTGAGGGGATCATATCACTAAGTGCATTTTCAAGTCTTCTGCTGAGCTT 1354
Db 256 GAGCAACGTTGTGAGGGGATCATATCACTAAGTGCATTTTCAAGTCTTCTGCTGAGCTT 197
Qy 1355 GTGAGGAGAGAGAGATGCGGAGTGTCTTTTAAACAGTAGCAACCTCTCGGCTTTGG 1414
Db 196 GTGAGGAGAGAGAGATGCGGAGTGTCTTTTAAACAGTAGCAACCTCTCGGCTTTGG 137
Qy 1415 GTCTTCGGATTAATGGCTATCTGAGGAGCTCAGGGTTCATGGTCTGCTGAGTTCCTT 1474
Db 136 GTCTTCGGATTAATGGCTATCTGAGGAGCTCAGGGTTCATGGTCTGCTGAGTTCCTT 85
Qy 1475 TTTTTCGTTTCGTTTTCATTTTACCATTTGGCCCTCCCGTGTGTTAAACAAATACATCAGGT 1534
Db 84 TTTTTCGTTTCGTTTTCATTTTACCATTTGGCCCTCCCGTGTGTTAAACAAATACATCAGGT 25
Qy 1535 GATTTTCATCTATTGGATTTGTTTC 1558
Db 24 GATTTTCATCTATTGGATTTGTTTC 1

RESULT 12

LOCUS CA211312 666 bp mRNA linear EST 25-SEP-2003
DEFINITION SCLAD1098B08.g Ad1 Saccharum officinarum cDNA clone SCLAD1098B08
CA211312
ACCSSION CA211312
VERSION CA211312.1 GI:35255493
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
REFERENCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 666)
Vettore A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCRST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
JOURNAL

COMMENT
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 098 row: B column: 08
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .666
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCLAD1098B08"
/lab_host="DH10B"
/clone_lib="Adi"
/note="Organ: seedlings inoculated with Gluconacetobacter
diazotrophicans; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[seedlings inoculated with Gluconacetobacter
diazotrophicans]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucet.lad.ic.unicamp.br/public"

FEATURES
source
Query Match 36.4%; Score 574; DB 14; Length 666;
Best Local Similarity 93.0%; Pred. No. 9e-101;
Matches 612; Conservative 0; Mismatches 45; Indels 1; Gaps 1;
QY 12 CATTGGCGATGCTCCTATCGAAGGCTGCCCACTGTATTTTCAGGCTATTTCAGTCTG 71
DB 1 CATTGGCGATGCTCCTATCGAAGGCTGCCCACTGTATTTTCAGGCTATTTCAGGCTG 60
QY 72 TCTCTGGGATTCACGAGTACAAATCTCTGACGGGGAATCGCGCATGATGAGCGT 131
DB 61 TCTCCAGGATCCACGAGTACAGATTTTGTGACGGGAGTGGCGCATGATGAGCGC 120
QY 132 CAACCTACCATATCTGGGAGTTTGGCATAGTTTAAACACACTTTACTTTCACAAAGGGAAT 191
DB 121 CAACCTACTATATCCGGGAGTTTGGTATAGTTTAAACACATTTGTACTTTCACAAAGGAAT 180
QY 192 AACCAATAACACCTTATCAAGTCCACGACACACTGGAAGCAGATGACATGATGTG 251
DB 181 AACCAATAACACCTTATTAAGTCCAAAGCAATTTGGAAGCAGATGAAATGATGTG 240
QY 252 GATTAATGAAATTTTCAACGTACCGTTACGTTGTTCAGATGCGACCGTTTCAGAGGACT 311
DB 241 GATTAATGAAATTTTCAACGTACCGTTACGTTGTTCGATGTACTGTTTCGAGGACT 300
QY 312 CTGAGATTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTGTTTCGAAATCTGAAT 371
DB 301 CCGAGAGTTTCAGAGGCTGCAATACAAATCTAGGTGTCGTGTTTCGAAATCTGAAT 360
QY 372 TTGCATACATGCTATGATTTTACTCCAGATTCGCGAGGTTATTCCTAGACATTAAT 431
DB 361 TTGCATACATGCTATGATTTTACTCCGATTCAGCGAGGTTATTCCTAGACATTAAT 420
QY 432 TTACTCTGAGGCAATCATTCATATTTCTCCATGAAACAGGGGATTCCTGTAGCTCCTC 491
DB 421 TTACTCTGAGGCAATCATTCATATTTCTCATGACAGGGGATTCCTGTAGCTCCTC 480
QY 492 TGGGACTCATTCAGAGGTCAAATTTGTTGGTCTCCTTAGCCCATTTGGAATTCATCTCAT 551
DB 481 TGGGACTCATTCAGAGGTCAAATTTGTTGGTCTCCTTAGCCCATTTGGAATTCATCTCAT 540

ORIGIN
Query Match 36.4%; Score 572.8; DB 14; Length 683;
Best Local Similarity 92.7%; Pred. No. 1.5e-100;
Matches 612; Conservative 0; Mismatches 47; Indels 1; Gaps 1;
QY 8 GGAACTTTGCGATGCTCTCTATCGAAGGCTGCCCACTGTATTTTCAGGCTATTTCAG 67
DB 25 GGAACTTTGCGATGCTCTCTATCGAAGGCTGCCCACTGTATTTTCAGGCTATTTCAG 83
QY 68 TCTGCTCTCTGGGATTCACGAGTACAAATCTTTGTCGACGGGAATGCGGCGCATGATGA 127
DB 84 CCTGCTCTCCAGGGATTCACGAGTACAAAGCCTTTTGTGACCGGGAGTGGCGGCGCATGATGA 143
QY 128 GCGTCAACCTACCATATCTCGGGAGTTTGGCATAGTTTAAACACACTTTTACTTTGACAGGGA 187

QY 552 TTGCGGAGCTAGAAACTCATGGCTCGAACTTGACAGAGAGCAGCTTGAAACACACACT 611
DB 541 TTGCGGAGCTAGAAACTCATGGCTCGAACTTGACAGAGAGCAGCTTGAAACACACACT 600
QY 612 ATATCTCATGGAAGAGGCTTAAGCGCAAC-ATAATGGAAGAAATGATAGTCAATGG 668
DB 601 ATATCTCATGGAAGAGGCTTAAGCGCAACCTTTATCGAAGAAATGATAGTCAATGG 658

RESULT 13
CA262684
LOCUS
DEFINITION
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CA262684
SCPIB2024A07.g LB2 Saccharum officinarum cDNA clone SCPIB2024A07
CA262684
CA262684.1 GI:35954887
EST.
Saccharum officinarum
Saccharum officinarum
REFERENCE
1 (bases 1 to 683)
Vettore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 024 row: A column: 07
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .683
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCPIB2024A07"
/lab_host="DH10B"
/clone_lib="LB2"
/note="Organ: Lateral buds from plants adult plants
growing in greenhouse; Vector: pSport1; Site 1: SalI;
Site 2: NotI; An unidirectional cDNA library generated
from [lateral buds from plants adult plants growing in
greenhouse]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucet.lad.ic.unicamp.br/public"

Db 144 GCGCAACCTACTATATCCGGGAGTTTGGTATAGTTAAACACATTGTACTTGAACAGAGA 203
Qy 188 ATATAACCAATAAACAACCTTATCAAGTCCAGACACACCTTGAAGCAGGATGACATGGA 247
Db 204 ATTAAACCAATAAACAACCTTATTAAGTCCAGACACAAATTTGAAGCAGGATGACATGGA 263
Qy 248 TGTGGAATGAATAATTTCAACCTAGCGTTACCTGTGTGATGCGACCGTTTCAGAAAGG 307
Db 264 TGTGGAATGAATAATTTCAACCTAGCGTTACCTGTGTGATGCGACCGTTTCAGAAAGG 323
Qy 308 TACTCTGAGATTTCAGAGGTCGAATACAAATATCTAGGTGTCGTGTTCTGAATATCT 367
Db 324 TACTCCGAGATTTCAGAGGTCGAATACAAATCTCTAGGTGCGCGTTTCTGAATATCT 383
Qy 368 GAATTTGCATACATGCTATGATTTACTCCAGATTCTGCGAGGTTATTGCCCTAGACNT 427
Db 384 GAATTTGCATACATGCTATGATTTACTCCGATTTCAGGCAAGGTTATTGCTCTGGAANT 443
Qy 428 TAATTTACCTGTGAAGCAATCAATCCCAATTTCTCCATGAACAGCGGATTCCTGTAGCTCC 487
Db 444 TAATTTACCTGTGAAGCAATCTTTTCAATTTCTACATGAACAGCGGATTCCTGTAGCTCC 503
Qy 488 TCTCTGGAATCTAATCAGAGTCAATTTGTTGGTCTCTTAGCCCATTTGGAATTTCAACT 547
Db 504 TCTCTGGAATCTAATCAGAGTCAATTTGTTGGTCTCTTAGCCCATTTGGAATTTCAACT 563
Qy 548 CATATTGCGGGAGCTAGAACTCATGCTCGAATCTGACAGACAGAGCAGCTTGAACACA 607
Db 564 TATATTGCGGGAGCTAGAACTCATGCTCGAATCTGACAGACAGAGCAGCTTGAACACA 623
Qy 608 CACTATATCTGCATGGAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATAGTCAGTG 667
Db 624 CACTATATCTGCATGGAAGAGGCTAAGCGGCAAACTAATGGAAGAAATGATAGTCAGTG 683

RESULT 14
CF011160 574 bp mRNA linear EST 17-JUL-2003
LOCUS QB36f03.xg QB3 Zea mays cDNA clone QB36f03, mRNA sequence.
DEFINITION CF011160
ACCESSION
VERSION
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 574)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (<http://www.genoplante.com>) and <http://genoplante-info.infobiogen.fr>.

FEATURES
source
1. .574
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="P2"
/db_xref="taxon:4577"
/clone="QB36f03"
/tissue_type="pollen"
/clone_lib="QB3"

ORIGIN
Query Watch 35.6%; Score 560.4; DB 14; Length 574;
Best Local Similarity 99.5%; Pred. No. 4.1e-98;
Matches 572; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 905 TCCGTCGGTTCCTGGGTTCCGAAATTTGGTATCTGAACAGTCTGTCATTCGGCTATGTT 964
Db 1 TCCGTCGGTTCCTGGGTTCCGAAATTTGGTATCTGAACAGTCTGTCATTCGGCTATGTT 60
Qy 965 GCGACCTAATAGCTTCACTTAGCTCTGCCCTTTAAACATGTTGGTTCAAGCTGGAGTAAGCTC 1024
Db 61 GCGACCTAATAGCTTCACTTAGCTCTGCCCTTTAAACATGTTGGTTCAAGCTGGAGTAAGCTC 120
Qy 1025 AATACCAATTTGGTATGATGACAAAGCTCCCTGCTTGAACATTTCTTGAAGTGAATCAC 1084
Db 121 AATACCAATTTGGTATGATGACAAAGCTCCCTGCTTGAACATTTCTTGAAGTGAATCAC 180
Qy 1085 AGCCTAGCAAAAGACAGAGTCTACACATGTTGGTGGATGAGATGACCATTCACA 1144
Db 181 AGCCTAGCAAAAGACAGAGTCTACACATGTTGGTGGATGAGATGACCATTCACA 240
Qy 1145 GGCCTTGACGCTTGGCAAGATGCAATACACCTTTTGGATTTCTTAACGCCAGAGATG 1204
Db 241 GGCCTTGACGCTTGGCAAGATGCAATACACCTTTTGGATTTCTTAACGCCAGAGATG 300
Qy 1205 CCAGATGCTCCGCTGCTGATCTCTTGGTGAAGTGGAGGAGCTGGCTAATCTCTGG 1264
Db 301 CCAGATGCTCCGCTGCTGATCTCTTGGTGAAGTGGAGGAGCTGGCTAATCTCTGG 360
Qy 1265 GGTGCGGGGTGTTTCAATTTGGAGCTGGAGGAGCTGGAGGAGCTGGCTAATCTCTGG 1324
Db 361 GGTGCGGGGTGTTTCAATTTGGAGCTGGAGGAGCTGGAGGAGCTGGCTAATCTCTGG 420
Qy 1325 AAGTGACATTTTCAAGTTCTTGTGCTGAGCTTGTGAGGAGAGGAGCATGGCGGATTTT 1384
Db 421 AAGTGACATTTTCAAGTTCTTGTGCTGAGCTTGTGAGGAGAGGAGCATGGCGGAG-TTTT 479
Qy 1385 TTTAAACAGTACCAACCTCTGGGCTTTGGGCTTCTGCGGATTAATGCGTATCGTCAGG 1444
Db 480 TTTAAACAGTACCAACCTCTGGGCTTTGGGCTTCTGCGGATTAATGCGTATCGTCAGG 539
Qy 1445 ACTCAGGTTTCATGCTGTAAGCTTCTCTTTT 1479
Db 540 ACTCAGGTTTCATGCTGTAAGCTTCTCTTTT 574

RESULT 15
BE123279 559 bp mRNA linear EST 14-JUN-2003
LOCUS 945040D06.Y1 945 - Mixed adult tissues from Walbot lab, same as 707
DEFINITION {SK} Zea mays cDNA, mRNA sequence.
ACCESSION BE123279
VERSION
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Sukarya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 559)
AUTHORS Walbot V.
TITLE Zea mays from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945040 row: D column: 06.

FEATURES
source
1. .559
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"

```

/db_xref="taxon:4577"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stages="fully-grown"
/lab_host="DH128"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/notes="organ: tassel, kernal, silk, husk, root, leaf;
vector: pGAD10; Site: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
issue ratio is 4:2:1:1:1 (tassel, kernal, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

```

ORIGIN

```

Query Match      35.4%; Score 557.4; DB 10; Length 559;
Best Local Similarity 99.8%; Pred. No. 1.6e-97;
Matches 558; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 197 AATAAACACCTTATCAAGTCCAGCACACCTGGAGCAGGATGACATGGATGGATAA 256
Db 1 AATAAACACCTTATCAAGTCCAGCACACCTGGAGCAGGATGACATGGATGGATAA 60

QY 257 TGAATAATTTCAACGTACCGTTACGTTGTGAGATGGACCGTTTCAGAAAGTACTCTGAG 316
Db 51 TGAATAATTTCAACGTACCGTTACGTTGTGAGATGGACCGTTTCAGAAAGTACTCTGAG 120

QY 317 AGTTTCAGAGGCTGCAATACAAATATCTAGGTTGTGTTTCTGAAATATCTGAAATTTGCA 376
Db 121 AGTTTCAGAGGCTGCAATACAAATATCTAGGTTGTGTTTCTGAAATATCTGAAATTTGCA 180

QY 377 CACATGCTATGATTTACTCCAGATTCCTGGCAGGTTATTGCCCTAGCACCTAAATTACC 436
Db 181 TACATGCTATGATTTACTCCAGATTCCTGGCAGGTTATTGCCCTAGCACCTAAATTACC 240

QY 437 TGTGAAGCAATCATTTCCATATTTCCATGAACAGGAGGATTCCTGTAGCTCCTCTCTGGGA 496
Db 241 TGTGAAGCAATCATTTCCATATTTCCATGAACAGGAGGATTCCTGTAGCTCCTCTCTGGGA 300

QY 497 CTCATTCAGAGGTCATTTGTTGGTCTCCTTAGCCGATTTGGATTTTCATATTTGGG 556
Db 301 CTCATTCAGAGGTCATTTGTTGGTCTCCTTAGCCGATTTGGATTTTCATATTTGGG 360

QY 557 GGAGCTAGAAACTCATGGCTCGAACTTGACAGAGCAGCTTTGAAACACACACTATATC 616
Db 361 GGAGCTAGAAACTCATGGCTCGAACTTGACAGAGCAGCTTTGAAACACACTATATC 420

QY 617 TGCATGGAAGAGGTTAAGCGGCAAACTAATCGAAGAAATGATAGTCACTGGCGACCGCA 676
Db 421 TGCATGGAAGAGGTTAAGCGGCAAACTAATCGAAGAAATGATAGTCACTGGCGACCGCA 480

QY 677 ACAGCATCTAGTGCATGGCCACCCCTTATGAGTCCCTTAGGGACATTCGAGTAAAGCTTTT 736
Db 481 ACAGCATCTAGTGCATGGCCACCCCTTATGAGTCCCTTAGGGACATTCGAGTAAAGCTTTT 540

QY 737 GCAAAATGGCATTTCCTACA 755
Db 541 GCAAAATGGCATTTCCTACA 559

```

Search completed: July 9, 2004, 08:14:41
Job time : 4377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: July 9, 2004, 04:35:33 ; Search time 777 Seconds
(without alignments)
9885.766 Million cell updates/sec

Title: US-09-857-525C-1

Perfect score: 1576

Sequence: 1 gcacagagggaacatttgcg.....tcaaaaaaaaaaaaaa 1576

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1423.4	90.3	1474	13	US-10-425-114-3817
2	1228	77.9	1844	13	US-10-425-114-1948
3	1226.4	77.8	2228	13	US-10-425-114-27250
4	1009	64.0	1387	13	US-10-425-114-33827
5	909	57.0	2980	17	US-10-437-963-2873
6	845.2	53.6	1200	13	US-10-425-114-6702
7	565.2	35.9	3077	13	US-10-424-599-44858
8	507	32.2	1507	13	US-10-425-114-29527
9	504.4	32.0	2315	13	US-10-424-599-79689
10	424.4	26.9	6747	17	US-10-437-963-9730
11	376	23.3	1112	17	US-10-437-963-22833
12	321.8	20.4	1225	13	US-10-424-599-101
13	310	19.7	721	13	US-10-425-114-30669
14	266.2	16.9	282	9	US-09-294-093B-135

15	241.6	15.3	1402	17	US-10-437-963-66100	Sequence 66100, A
16	219	13.9	1879	13	US-10-424-599-105	Sequence 105, App
17	217.6	13.8	1127	13	US-10-424-599-100	Sequence 100, App
18	188	11.9	1533	13	US-10-425-114-8857	Sequence 8857, App
19	176	11.2	1001	13	US-10-425-114-10694	Sequence 10694, A
20	175.2	11.1	365	12	US-09-732-627A-260	Sequence 260, App
21	163.2	10.4	954	13	US-10-425-114-2008	Sequence 2008, App
22	139	8.8	178	9	US-09-294-093B-4778	Sequence 4778, App
23	129.2	8.2	1140	13	US-10-424-599-44867	Sequence 44867, A
24	128.2	8.1	2362	13	US-10-425-114-31135	Sequence 31135, A
25	111.2	7.1	361	13	US-10-424-599-5459	Sequence 5459, App
26	100.2	6.4	411	13	US-10-424-599-44857	Sequence 44857, A
27	63.6	4.0	175	13	US-10-424-599-67266	Sequence 67266, A
28	58.8	3.7	373	13	US-10-424-599-55071	Sequence 55071, A
29	56.2	3.6	561	13	US-10-424-599-94913	Sequence 94913, A
30	54.4	3.5	917	13	US-10-424-599-44865	Sequence 44865, A
31	51.6	3.3	538	13	US-10-424-599-82474	Sequence 82474, A
32	50.6	3.2	11527	14	US-10-108-605-70	Sequence 70, Appl
33	47.2	3.0	517	17	US-10-021-323-11054	Sequence 11054, A
34	46.8	3.0	472	10	US-09-918-995-23770	Sequence 23770, A
35	45.6	2.9	485	17	US-10-021-323-6802	Sequence 6802, App
36	44.6	2.8	415	17	US-10-021-323-41	Sequence 41, Appl
37	44	2.8	200	17	US-10-021-323-16749	Sequence 16749, A
38	42.6	2.7	324	16	US-10-125-968-1261	Sequence 1261, App
39	42.4	2.7	1578	16	US-10-305-720-1113	Sequence 1113, App
40	42.4	2.7	1691	9	US-09-925-297-2	Sequence 2, Appl
41	42.2	2.7	391	13	US-10-424-599-59659	Sequence 59659, A
42	42.2	2.7	512	10	US-09-814-353-5072	Sequence 5072, App
43	42.2	2.7	512	10	US-09-814-353-11364	Sequence 11364, A
44	42	2.7	499	17	US-10-021-323-10189	Sequence 10189, A
45	42	2.7	1435	10	US-09-763-970-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-10-425-114-3817

; Sequence 3817, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 3817

; LENGTH: 1474

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700342215_FLI

US-10-425-114-3817

Query Match 90.3%; Score 1423.4; DB 13; Length 1474;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1457; Conservative 0; Mismatches 6; Indels 10; Gaps 2;

QY 86 CGAGTACAAATCTTTGTGGACGGGAATGGGGGATGATGAGCGTCAACCTACCATATC 145

Db 1 CGAGTACAAATCTTTGTGGACGGGAATGGGGGATGATGAGCGTCAACCTACCATATC 60

QY 146 TGGGAGATTGGCATAGTTAACACACTTTACTTGACAAGGGAATATAACCAATAAACAC 205

Db 61 TGGGAGATTGGCATAGTTAACACACTTTACTTGACAGGGGAATATAACCAATAAACAC 120

206 CTTATCAAGTCCAAAGCACACCTCGAGACAGGATGACATGATGTGGATTAATGAATTTT 265
 121 CTTATCAAGTCCAAAGCACACCTCGAGACAGGATGACATGATGTGGATTAATGAATTTT 180
 266 TCAACCTTACGCTTACGCTTGTGACATGACACCGTTTCAGAGGTACTCTGAGAGTTTCAGA 325
 181 TCAACCTTACGCTTACGCTTGTGACATGACACCGTTTCAGAGGTACTCTGAGAGTTTCAGA 240
 326 GCGTCAATACAAATATCTAGGTGTGCTGTTTCGAAATATCTGAATTTGCAATGCTA 385
 241 GCGTCAATACAAATATCTAGGTGTGCTGTTTCGAAATATCTGAATTTGCAATGCTA 300
 386 TGAATTTACTCCAGATCTCGAGAGGTATGTCCTGACATTAATTTACTCTGAGACA 445
 301 TGAATTTACTCCAGATCTCGAGAGGTATGTCCTGACATTAATTTACTCTGAGACA 360
 446 ATCAATTCATATCTCCATGAACAGGGGATTCCTGTAGCTCCCTCTGAGACTCAATCAG 505
 361 ATCAATTCATATCTCCATGAACAGGGGATTCCTGTAGCTCCCTCTGAGACTCAATCAG 420
 506 AGGTCAATTTGGTGTCTCTTAGCCCAATTTGATTTTCATATCTCATATTTGCGGAGCTAGA 565
 421 AGGTCAATTTGGTGTCTCTTAGCCCAATTTGATTTTCATATCTCATATTTGCGGAGCTAGA 480
 566 AACTCATGCTCGAATCTGACAGAGAGAGCTTGAACACACATATATCTCATGAGAA 625
 481 AACTCATGCTCGAATCTGACAGAGAGAGCTTGAACACACATATATCTCATGAGAA 540
 626 AGAGGCTAAGCGGCAAACTAATGGAAGAAATGATGATGAGTCACTGCGAGCGCAACAGCATCT 685
 541 AGAGGCTAAGCGGCAAACTAATGGAAGAAATGATGATGAGTCACTGCGAGCGCAACAGCATCT 600
 686 AGTGAATGCCACCCCTATGAGTCTCTGAGGACATTTGAGTAAGCTTTTGAATAGG 745
 601 AGTGAATGCCACCCCTATGAGTCTCTGAGGACATTTGAGTAAGCTTTTGAATAGG 660
 746 CATTTCTACAGTCCAGTATTTATTTATCATCATCATCATGATGAGTCACTTCCCGAGTTATT 805
 561 CATTTCTACAGTCCAGTATTTATTTATCATCATCATCATGATGAGTCACTTCCCGAGTTATT 720
 806 GCATCTTGCAATCTTTCTGGAATTTTGAATGATTTGATGATGATTTGATGATGATTTCAAC 865
 721 GCATCTTGCAATCTTTCTGGAATTTTGAATGATTTGATGATGATTTGATGATGATTTCAAC 780
 866 TGGTAATTTGCCATTTCTGACCAACCAAGTGTGCTCCATTCGGTTCCTGGGTTC 925
 781 TGGTAATTTGCCATTTCTGACCAACCAAGTGTGCTCCATTCGGTTCCTGGGTTC 840
 926 GAAATTTGGTGTATCTGAACAGTCTCCATTTGGTGTATGTTGCGACCTTAATGCTCACTTAG 985
 841 GAAATTTGGTGTATCTGAACAGTCTCCATTTGGTGTATGTTGCGACCTTAATGCTCACTTAG 900
 986 CTCTGCGCTTAAATGTTGTTTCAAGCTGAGTAAGCTCAATACCAATTTGGATGACAA 1045
 901 CTCTGCGCTTAAATGTTGTTTCAAGCTGAGTAAGCTCAATACCAATTTGGATGACAA 960
 1046 CGACTCCCTGTTGACATTTCTAGAGTGAATCATCAGACCTTACGAAAGACAGAGT 1105
 961 CGACTCCCTGTTGACATTTCTAGAGTGAATCATCAGACCTTACGAAAGACAGAGT 1020
 1106 CTACACATGTTTCCGCTGATGAGATGACATTTCAACAGGCTTTGAGCTTGGACAAGA 1165
 1021 CTACACATGTTTCCGCTGATGAGATGACATTTCAACAGGCTTTGAGCTTGGACAAGA 1080
 1166 TGCCAATACACCTTTTGGATTTCTTAACGCGCAGAGATGCGAGATGTCCTCGGTCTGA 1225
 1081 TGCCAATACACCTTTTGGATTTCTTAACGCGCAGAGATGCGAGATGTCCTCGGTCTGA 1140
 1226 TCCCTTTGCTGAAGTGTGATGAGGACCTGGCTAATCTCGGGTGGCGGGGTTCATTTGT 1285
 1141 TCCCTTTGCTGAAGTGTGATGAGGACCTGGCTAATCTCGGGTGGCGGGGTTCATTTGT 1200
 1286 GGAAGCTGGGAGCAACCTGTGGAGGCGCATCATCACTAAGTGAATTTTCAAGTTCTT 1345

1201 GGAAGCTGGGAGCAACCTGTGTGAGGGCATCATATCACTAAGTACATTTTCAAGTTCTT 1260
 1346 GCTGAGCTTCTGAGGAGAGAGGAGCATGCGGAGTGTGTTTTTAACTAGTAGCAACCTCTG 1405
 1261 GCTGAGCTTCTGAGGAGAGAGGAGCATGCGGAGTGTGTTTTTAACTAGTAGCAACCTCTG 1320
 1406 GGCCTTTGGTCCCTGCGGATAATTTGGCTATGTCAGGACTCAGGGTTCATGCTGTGA 1465
 1321 GGCCTTTGGTCCCTGCGGATAATTTGGCTATGTCAGGACTCAGGGTTCATGCTGTGA 1372
 1466 AAGTTCCTTTTTCCTGTTTTTACATTTGGCTCCCTCCGTTGTTGTTAAACAAACAT 1525
 1373 AAGTTCCTTTTTCCTGTTTTTACATTTGGCTCCCTCCGTTGTTGTTAAACAAACAT 1430
 1526 ACATCAGGTGATTTTATCTATTTGGATTGTTTC 1558
 1431 ACATCAGGTGATTTTATTTATTTGATTGTTTC 1463

RESULT 2

US-10-425-114-1948
 ; Sequence 1948, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 1948
 ; LENGTH: 1844
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700172325_FLI
 US-10-425-114-1948

Query Match 77.9%; Score 1228; DB 13; Length 1844;
 Best Local Similarity 93.8%; Pred. No. 0;
 Matches 1279; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 8 GGAACATTTGCCGATGTCCTCTATGAGGCTGCCACCTGATTTTCAGGCTATTTGCAG 67
 DB 205 GGAACATTTGCCGATGTCCTCTGAGAGGCTGCCACCTGATTTTCAGGCTATTTGCAG 264
 QY 68 TCTGTCTCTCGGATTCACGAGTACAAATTTCTTTGTGACGGGAAATGGCGGATGATGA 127
 DB 265 CTTGTCTCCAGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 324
 QY 128 GCTGACCTTACCATATCTGGGAGTGTGCTATGATGATGATGATGATGATGATGATGATGATGA 187
 DB 325 GCGCCAACTTACTATATCTGGGAGTTCGGTATGATGATGATGATGATGATGATGATGATGATGA 384
 QY 188 ATATAACCAATTAACACACCTTATCAAGTCCAAAGCACACCTTGAAGAGGATGAACATGGA 247
 DB 385 ATTATCAATTAATTAATGCTTATTAATCAAGCACACCTTGAAGAGGATGAACATGGA 444
 QY 248 TGTGATAATGAATTTTCAACGTCAGGTTACGTTGTGATGATGATGATGATGATGATGATGATGATGAAGG 307
 DB 445 TGTGATAATGAATTTTCAACATACGTTTACATTTGTGATGATGATGATGATGATGATGATGATGATGAAGG 504
 QY 308 TACTCTGAGGTTTCAAGGCTGCAATCAAAATATCTAGGTTGCTGTTTCTGATATCT 367
 DB 505 TACTGTGAGGTTTCAAGGCTGCAATCAAAATCTCAAGGTGCGGCTTCTGATATCT 564

368 GAATTTCATACATGCTATGATTTACTCCAGATTCCTGGCAAGGTTATTGCCCTAGACAT 427
Db
565 CAATTTCATACATGCTATGATTTACTCCAGATTCCTGGCAAGGTTATTGCCCTAGACAT 624
Qy 428 TAATTTCATGGAAGCAATCAATTCATATTTCTCCATGAAAGGGAATTCCTAGCTCC 487
Db 625 TAATTTCATGGAAGCAATTCCTTCATATTTCTGCATGAAAGGGAATTCCTAGCTCC 684
Qy 488 TCTCTGGGACTCATTCAGAGGCTCAATTTGTGGTCTCCTTAGCCCATTTGGATTTCATCT 547
Db 685 TCTCTGGGACTCATTCAGAGGCTCAATTTGTGGTCTCCTTAGCCCATTTGGATTTCATCT 744
Qy 548 CATATTTCGGGAGCTAGAACTCATGGCTCGAATTTGACAGAGAGAGAGAGAGAGAGAGAG 607
Db 745 TATATTTCGGGAGCTAGAACTCATGGCTCGAATTTGACAGAGAGAGAGAGAGAGAGAGAG 804
Qy 608 CACTATATCTGCATGAAAGAGAGCTAAGCGGCAAACTTAATGGAAGAAATGATGTCAGTG 667
Db 805 CACTATATCTGCATGAAAGAGAGCTAAGCGGCAAACTTTGTGGAAGAAATGATGTCAGTG 864
Qy 668 GCGACCGCAACAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTGAGGACATTCGAGT 727
Db 865 GCGAGCAGCATCAGCCTGATGATGCGCCCTTATGAGTCTCTGAGGACATTCGAGT 924
Qy 728 AAAGCTTTTGCAGAAATGGCAATTTCTACAGTGCAGTTTATTTATTCATCATCATCATGG 787
Db 925 AAAGCTTTTGTCTAAATGACATTTCTACAGTGCAGTTTATTTATTCATCATCATCATGG 984
Qy 788 ATCAATTCGGCAGCTTATGATCTTCGATCAGTCTTCGAAATTTGAAATGATTTGTAG 847
Db 985 ATCAATTCGGCAGCTTATGATCTTCGATCAGTCTTCGAAATTTGAAATGATTTGTAG 1044
Qy 848 ATACTTCAAAATCAACTGTAATTTGCTATTCTGAAACCAACAGTGTGCTCCATTC 907
Db 1045 GTATTTTAAATCTCACTGTAATTTGCTATTCTGAAACCAACCGGTGTGCTCCATTC 1104
Qy 908 GCTGGGTTCTCGGTTCCGAAATTTGGTGTATCTGAAACAGTGTGCTCAATTCGTTGG 967
Db 1105 GCTGGGTTCTCGGTTCCGAAATTCGGTGTATCTGAAACAGTGTGCTCAATTCGTTGG 1164
Qy 968 ACCTAATGCTCAGTATGCTGCTGCTTAAATGATGCTGCTGCTGCTGCTGCTGCTCAAT 1027
Db 1165 ACCTAATGCTCAGTATGCTGCTGCTTAAATGATGCTGCTGCTGCTGCTGCTGCTCAAT 1224
Qy 1028 ACCAATGTGATGACAAAGCTCCCTGCTGACACTTACTCTAGAAAGTGCATCAGC 1087
Db 1225 ACCAATGTGATGACAAAGCTCCCTGCTGACACTTACTCTAGAAAGTGCATCAGC 1284
Qy 1088 CCTAGCAAAAGCAAGGCTTACACATGTTGGGCTGATGAGATGACCAATTCACGAGC 1147
Db 1285 CCTAGCTAAAGCAAGGCTTACACATGTTGGGCTGATGAGATGACCAATTCACGAGC 1344
Qy 1148 TTTGAGCTTGGACAGATGCCAATACACTTTTGGATTCTTTAAACCGGCGAGAGTSCCA 1207
Db 1345 TTTACAGCTTGGACAGATGCCAATACACTTTTGGATTCTTTAAACCGGCGAGAGTSCCA 1404
Qy 1208 GATGTGCTCCGCTCTGATCCTTTGCTGAAGGTTGATGAGAGCACTGGCTAATCTCGGGT 1267
Db 1405 GATGTGCTCCGCTCTGATCCTTTGCTGAAGGTTGATGAGAGCACTGGCTAATCTCGGGT 1464
Qy 1268 GCGCGGGGTTCAATGTGAAAGCTGGAGCAAACTGTGTGAGGGGATCATATCCTAAG 1327
Db 1465 GCGCGGGGTTCAATGTGAAAGCTGGAGCAAACTGTGTGAGGGGATCATATCCTAAG 1524
Qy 1328 TGACATTTCAAGTTCTTCTGAGCTTCTGAGGAGAGAGAGGCA 1371
Db 1525 TGATATTTCAAGTTCTTCTGAGCTTCTGAGGAGAGAGAGGCA 1568

RESULT 3

US-10-425-114-27250

; Sequence 27250, Application US/10425114

; Publication No. US2004003488A1

GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27250
; LENGTH: 2228
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: L1B4729-018-A12_FLI
US-10-425-114-27250

Query Match 77.8%; Score 1226.4; DB 13; Length 2228;

Best Local Similarity 93.7%; Pred. No. 0; Mismatches 86; Indels 0; Gaps 0;

Matches 1278; Conservative 0;

Qy 8 GGAACATTTCCGATGCTCTCTATCGAAGGCTGCCCACTGTATTTCAGGCTATTTCGAG 67
Db 588 GGAACATTTCCGATGCTCTCTATCGAAGGCTGCCCACTGTATTTCAGGCTATTTCGAG 647
Qy 68 TCTGTCTCTGGGATTCACAGTACAAATTTCTTTGTGACGGGAATTCGGGCGATGATGA 127
Db 648 CTTGTCTCCAGGATTCATGAGTACAAATTTCTATGTGACGGGAGTGGCGTCATGATGA 707
Qy 128 GCGTCAACCTACCATATCTCGGGAGTTTGGCATAGTTAAACACACTTTACTTGAACAAGGA 187
Db 708 GCGCAACCTTACTATATCTCGGGAGTTTGGCATAGTTAAACACACTTTACTTGAACAAGGA 767
Qy 188 ATATAACCAATAAACAACCTTATCAAGTCCCAAGCACACCTGGAAGCAGGATGAACATGA 247
Db 768 ATTATCAATAAATGCTTTTAAATCCCAAGCACACCTGGAAGCAGGATGAACATGA 827
Qy 248 TGTGGATTAATGAAATTTTCAAGTACGTTTACGTTTGTAGATGCGACCGTTTCGAAGG 307
Db 828 TGTGGATTAATGAAATTTTCAAGTACGTTTACGTTTGTAGATGCGACCGTTTCGAAGG 887
Qy 308 TACTCTCAGAGTTTTCAGAGCTGCAATACAAATATCTAGGTGTCGTGTTCTGAAATATCT 367
Db 888 TACTCTCAGAGTTTTCAGAGCTGCAATACAAATATCTAGGTGTCGTGTTCTGAAATATCT 947
Qy 368 GAATTTGCATACATGCTATGATTTTATCCAGATTCCTGGCAAGSTTATTCGCTAGACAT 427
Db 948 CAATTTGCATACATGCTATGATTTTATCCGATTCGCGAAGTTCGCTAGACAT 1007
Qy 428 TAATTTACCTGTGAAGCAATTCATATCTCCATATCTCCATGACAGCGGATTCCTGTAGCTCC 487
Db 1008 TAATTTACCTGTGAAGCAATTCATATCTCCATGACAGCGGATTCCTGTAGCTCC 1067
Qy 488 TCTCTGGGACTCATTCAGAGTCAATTTGTGGTCTCCTTAGCCCATTTGGATTTCATCT 547
Db 1068 TCTCTGGGACTCATTCAGAGTCAATTTGTGGTCTCCTTAGCCCATTTGGATTTCATCT 1127
Qy 548 CATATTTCGGGAGCTAGAACTCATGCTCGAATTCGAACTTGCACAGAGAGAGAGAGAGAG 607
Db 1128 TATATTTCGGGAGCTAGAACTCATGCTCGAATTCGAACTTGCACAGAGAGAGAGAGAGAG 1187
Qy 608 CACTATATCTGCATGAAAGAGGCTAAGCGGCAAACTTAATGGAAGAAATGATGTCAGTG 667
Db 1188 CACTATATCTGCATGAAAGAGGCTAAGCGGCAAACTTTGTGGAAGAAATGATGTCAGTG 1247
Qy 668 GCGACCGCAACAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTGAGGACATTCGAGT 727
Db 1248 GCGAGCAGCATCAGCATCTAGTGCATGCCACCCCTTATGAGTCTCTGAGGACATTCGAGT 1307

```
QY 728 AAAGCTTTTGCAGAAATGGCAATTTCTACAGTCCAGTATTTATTTATCATCATCAGATGG 787
Db 1308 AAAGCTTTTGCAGAAATGGCAATTTCTACAGTCCAGTATTTATTTATCATCATCAGATGG 1367
QY 788 ATCAATCCCGAGCTATGATCATCTTGGCATCACTTTCTGGAATTTTGAATGTAATTTCTAG 847
Db 1368 ATCAATCCCGAGCTATGATCATCTTGGCATCACTTTCTGGAATTTTGAATGTAATTTCTAG 1427
QY 848 ATACTTCAAAATCAAACTGGTAAATTTGGCTATTTCTGAACCAACAGCTGGTCCATTC 907
Db 1428 GTATTTTAAATCTCAATGGTAAATTTGGCTATTTCTGAACCAACAGCTGGTCCATTC 1487
QY 908 GCTGGGTTCTGGGTTCCGAAATTTGGTGAATCTGAACAGTCTGCATTTGGTGAATTTGGG 967
Db 1488 GCTGGGTTCTGGGTTCCGAAATTTGGTGAATCTGAACAGTCTGCATTTGGTGAATTTGGG 1547
QY 968 ACCTAATGCCCTCACTAGCTCTGGCTTAAACATGTTGGTTCAGCTGGATGAAGCTCAAT 1027
Db 1548 ACCTAATGCCCTCACTAGCTCTGGCTTAAACATGTTGGTTCAGCTGGATGAAGCTCAAT 1607
QY 1028 ACCAATTTGGATGACCAACAGCTCCCTGCTTGACACTTTACTCTAGAAAGTGACATCAGC 1087
Db 1608 ACCAATTTGGATGACCAACAGCTCCCTGCTTGACACTTTACTCTAGAAAGTGACATCAGC 1667
QY 1088 CCTAGAAAAGAGAGGTCTACACATGTTTGGCTGGATGAGATGACATTCACAGGC 1147
Db 1668 CCTAGAAAAGAGAGGTCTACACATGTTTGGCTGGATGAGATGACATTCACAGGC 1727
QY 1148 TTTCAGCTTGGCAAGATGCCAATACACCTTTTGGATTTCTTTAAAGCCAGAGATGCCA 1207
Db 1728 TTTCAGCTTGGCAAGATGCCAATACACCTTTTGGATTTCTTTAAAGCCAGAGATGCCA 1787
QY 1208 GATGTGCTCCGGTCTGATCCTTTGCTGAAGGTGATGAGAGCAGCTGGCTAATCTCGGGT 1267
Db 1788 GATGTGCTCCGGTCTGATCCTTTGCTGAAGGTGATGAGAGCAGCTGGCTAATCTCGGGT 1847
QY 1268 GCGCGGGTGTTCATTTGGAGGTGGAGGCAACAGTGTGGAGGCATCATCATCTAAG 1327
Db 1848 GCGCGGGTGTTCATTTGGAGGTGGAGGCAACAGTGTGGAGGCATCATCATCTAAG 1907
QY 1328 TGCAATTTCAAGTTCTTGGTGAAGTGTGAGGAGAGAGGAGCA 1371
Db 1908 TGCAATTTCAAGTTCTTGGTGAAGTGTGAGGAGAGAGGAGCA 1951
```

RESULT 4

```
US-10-425-114-33827
; Sequence 33827, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33827
; LENGTH: 1387
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLM017174G09_FLI
US-10-425-114-33827
```

```
Query Match 64.0%; Score 1009; DB 13; Length 1387;
Best Local Similarity 94.2%; Pred. No. 8.5e-296;
Matches 1048; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
```

```
QY 259 AAAATTTTCAACGTAACGGTTAAGTTGTCAGATGSCACGGTTTCAAGAGTACTCTGAGAG 318
Db 1 AAAATTTTCAACGTAACGGTTAAGTTGTCAGATGSCACGGTTTCAAGAGTACTCTGAGAG 60
QY 319 TTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTGTTTCTGAAATATCTGAATTTGCAATA 378
Db 61 TTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTGTTTCTGAAATATCTGAATTTGCAATA 120
QY 379 CATGCTATGATTTTACTCCAGATTTCTGGCAAGGTTATTTGCCCTAGACATTAATTTTACCTG 438
Db 121 CATGCTATGATTTTACTCCGATTTCCGGCAAGGTTATTTGCTCTCGACATTAATTTTACCTG 180
QY 439 TGAAGCAATCATTCATATTTCTCATGAACAGGGGATTCCTGTAGTCTCTCTCGGACT 498
Db 181 TGAAGCAATCTTTTTCATATTTCTGCATGAACAGGGGATTCCTGTAGTCTCTCTCGGACT 240
QY 499 CATTCAGAGGTCATTTGTTGTTGTTCTCTAGCCATTTGGATTTTCATCTCATATTTCCGGG 558
Db 241 CATTCAGAGGTCATTTGTTGTTGTTCTCTAGCCATTTGGATTTTCATCTCATATTTCCGGG 300
QY 559 AGCTAGAAACTCATGCTCGAACTTTGAACAGAGGAGCTTTGAAACACACTATATCTG 618
Db 301 AGCTAGAAACTCATGCTCGAACTTTGAACAGAGGATCAGCTTGAACACACTATATCTG 360
QY 619 CATGGAAGAGGCTAAGGGGCAAACTAATGGAAGAAATGATAGTCAGTGGGAGCCGCAAC 678
Db 361 CATGGAAGAGGCTAAGGGGCAAACTTGTGGAAGAAATGATAGTTCAGTGGGAGGACATC 420
QY 679 AGCATCTAGTGCATGCAACCCCTTATGAGTCTCTTGAGGAGCATTTGCAAGCTTTTGC 738
Db 421 AGCATCTAGTGCATGCAACCCCTTATGAGTCTCTTGAGGAGCATTTGCAAGCTTTTGC 480
QY 739 AAAATGGCAATTTCAAGTGCAGTATTTATTCATCATCATCATCATCATCATCATCATCAT 798
Db 481 AAAATGCAATTTCTACAGTGCAGTATTTATTCATCATCATCATCATCATCATCATCATCAT 540
QY 799 AGTTATTTGCACTTTGCACTCACTTTCTGGAATTTTGAATGATTTTGTAGATTAATTTCAAAA 858
Db 541 AGTTATTTGCACTTTGCACTCACTTTCTGGAATTTTGAATGATTTTGTAGATTAATTTCAAAA 600
QY 859 ACTCAACTGGTAAATTTGCTTATTTCTGAACCAACAGTGTGCTTCCATTCCTCGGTGGTCT 918
Db 601 ACTCAACTGGTAAATTTGCTTATTTCTGAACCAACAGTGTGCTTCCATTCCTCGGTGGTCT 660
QY 919 GGGTTCCGAAAAATTTGGTGCATCTGAAACAGTGTGCTTCCATTTGTTGGAATTAATTCCT 978
Db 661 GGGTTCCGAAAAATTTGGTGCATCTGAAACAGTGTGCTTCCATTTGTTGGAATTAATTCCT 720
QY 979 CACTTAGCTCTGCTTAAACATTTGTTTCAAGCTGGAAGTGAAGTCAATACCAATTTGGG 1038
Db 721 CACTTAGCTCTGCTTAAACATTTGTTTCAAGCTGGAAGTGAAGTCAATACCAATTTGGG 780
QY 1039 ATGACAAAGCACTCCCTGCTTGACACTTACTCTAGAAGTGACATCAAGCCCTAGCAAAAG 1098
Db 781 ATGAAAACGACTCCCTGCTTGACACTTACTCTAGAAGTGACATCAAGCCCTAGCAAAAG 840
QY 1099 ACAAGTCTTACACATGTTGGCTGGATGAGATGACCATTCACAGGCTTTGCAAGCTTG 1158
Db 841 ACAAGGCTTACACATGTTGGCTGGATGAGATGACCATTCATCAGGCTTTTACAGCTTG 900
QY 1159 GACAAGATGCCAATACACTTTTGGATTTCTTAAACGCGCAGAGATGCCAGATGTCCTCC 1218
Db 901 GACAAGATGCCAATACACTTTTGGATTTTAAACGCTCAGAGATGCCAGATGTCCTCC 960
QY 1219 GGTCTGATCTTTGCTGAAGGTGATGGAGCACTGGCTTAATCTCGGGGTGGGCGGGTGT 1278
Db 961 GGTCTGATCTTTGCTGAAGGTGATGGAGCACTGGCTTAATCTCGGGGTGGGCGGGTGT 1020
QY 1279 TCATTTGGGAAGCTGGGAGGAAACGTTGGAGGGCATTCATATCACTAAGTGAATTTTCA 1338
Db 1021 TCATAGTAGAAGCTGGGAGGAAACGTTGGAGGGTATTTATATCACTAGTGTATTTTCA 1080
```

QY 1339 AGTCTTCTGCTGAGCTTGTGAGGAGAGAGCA 1371
|||||
Db 1081 AGTCTTCTGAGCTTGTGATAGAGAGGGA 1113
|||||
RESULT 5
US-10-437-963-2873
; Sequence 2873, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 2873
; LENGTH: 2980
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT453c_102608C.1
US-10-437-963-2873
Query Match 57.7%; Score 903; DB 17; Length 2980;
Best Local Similarity 82.6%; Pred. No. 4.1e-265;
Matches 1041; Conservative 0; Mismatches 220; Indels 0; Gaps 0;
QY 5 GAGGGAACATTTGCCGATGCTCTCTATCGAAGGCTGCCCACTGTATTTTCAGGCTATTG 64
|||||
Db 681 GACAGACATTTGCCAATGCTCTCTGTGAAGTGTGCCACCGTTTTCAGGCCATATG 740
|||||
QY 65 GAGTGTCTCTGGGATTCAGAGTACAAATCTTTGTGGACGGGATCGGGGATGA 124
|||||
Db 741 CAGTTTGTCCCTGGGATTCATCAGTACAAATTTTGTGTGGATGGGGAATGGCGCACGA 800
|||||
QY 125 TCAGGTCACACTACCATATCTGGGAGTTTGGCATAGTTTAAACACACTTTACTTTGACAAG 184
|||||
Db 801 TGAGGGCAACCTACCTAACAGGAGACTATGGTGTGGTTTAACTTTTGTGCTAGTAG 860
|||||
QY 185 GGAATATAACCAATAAACACCTTATCAAGTCCAGCACACCTGGGAAGCAGGATGAACAT 244
|||||
Db 861 GGAATTTGACAGATAAATAACAATATTGAGCCCTAGTACACCTGGGAAGTAGGATGAACAT 920
|||||
QY 245 GGAATGGATTAATAAATTTCAAGTAGGTTAGTTGTGATGGGACCGTTTCAGA 304
|||||
Db 921 GGAATGGACAAGCAATTTTCAAGTACTGTCTTTTGTCTGATGGCATTTATTCAGGA 980
|||||
QY 305 AGGTACTCTGAGATTTTCAGAGCTCCAAATACAAATATCTAGTGTGCTGTTTCTCAATA 364
|||||
Db 981 AGGTCTCAGAGATTTTCAGAGCAGCTATACAGATCTCTAGTGTGCTGTTGAGCGATTT 1040
|||||
QY 365 TCTGAATTTGCATACATGCTATGATTTACTTCCAGATTTCTGGCAAGGTTATTTGCCCTAGA 424
|||||
Db 1041 TCTGAATGACAACTGGGTATGATTTACTTCCAGATTTCTGGCAAGGTCATTTGCTCTAGA 1100
|||||
QY 425 CATTAATTTACCTGTGAAGCAATCATCTATCTCCATGACAGGGAATTCCTGTAGC 484
|||||
Db 1101 CCAATTTTGCCTGTGAAGCAATCTTTTCAATTTCTTCAATGACAGGGAATTCCTGTGGC 1160
|||||
QY 485 TCTCTCTGGGACTCATTCAGAGTCAATTTGTGTGCTCTCTTACGCCATTTGATTTCTAT 544
|||||
Db 1161 ACCTCTATGGGATTCATTCAGGGGCGAGTTTGTGGCCCTTTTGTAGCCCACTGTGATTTAT 1220
|||||
QY 545 ACTCATATTGGGGAGCTAGAAATCTATGCTCGAATTTGACAGAGAGAGCGCTTGAAC 604
|||||

Db 1221 ACTCATATTGAGAGAGCTGAAACTCATGCTCTCAATCTGACAGAGAGAGCGTTGAAC 1280
|||||
QY 605 ACACACTATATCTGCATGGAAGAGGCTAAGCGGAACTAAATGGAGAAATGATATCTCA 664
|||||
Db 1281 ACATACATATCTGCTGGAAGAGGCGCAGCGTCAAACTTATGCCAAGAACGAGGCTTC 1340
|||||
QY 665 GTGGCGACCGCAACAGCATCTAGTCATGSCACCCCTTATGAGTCTCTGAGGACATTCG 724
|||||
Db 1341 TTGGAGGGCAATCACCATTAGTTTATGSCACCCCTTATGATCCCTGAGGGAATTCG 1400
|||||
QY 725 AGTAAAGCTTTTGCAAAATGGCATTCTACAGTGGCAGTTATTTATCATCATCAGAGA 784
|||||
Db 1401 TATGAAGATACCTGCAAAACGGTGTCTACCGTTCCAAATATGTTTCTCCATCATCACCAGA 1460
|||||
QY 785 TGGATCATTTCCCGCAGTTATTCATCTGCATCCTGATCCTCTGGAATTTGAAATGATTTG 844
|||||
Db 1461 TGCTCATATCCCAATTTGTCATCTGCTCCCTTCTGGAATTTGAAATGATTTG 1520
|||||
QY 845 TAGATACCTTCAAAATCTCAACTGTAATTTGCTTATTTCTGAAACCAACAGTGTCTCCAT 904
|||||
Db 1521 TAGATATTTTAAATTTCTCAAGTAAATTTACCTATTTTGGCCACACTCTCATGCCAAT 1580
|||||
QY 905 TCCGCTGGGTTCTTGGGTTCCGAAATTTGGTGAATCTGAACAGTCTGTCATTTGCTATGTT 964
|||||
Db 1581 TCCTCTGGGTACCTGGGTTCCAAAATTTGGTGAATCTCTATGCGCTCCATTTGGCTATGTT 1640
|||||
QY 965 GGGACCTAATGCTCTCATTAGCTCTGCCCTTACACATGTTGGTTCAAGCTGGAGTAAAGCTC 1024
|||||
Db 1641 GCGGCTTACACATCTCTTAGCGCTGCCCTGAAATTTGCTGGTTCAAGCTGGTGGAGCTC 1700
|||||
QY 1025 AATPACCAATTTGGATGACAAAGCTCTCTGCTTGTGACACTTACTCTAGAAGTGAATCAC 1084
|||||
Db 1701 AATACCAATTTGGATGACAAAGCTCTGCTTGTGACACTTACTCTAGAAGTGAATCAC 1760
|||||
QY 1085 AGCCTTAGCAAAAGACAAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACCA 1144
|||||
Db 1761 AGCTCTTGCAAAAGACAAGGCTTACACATGTTGGCTGGATGAGATGACCATTCACCA 1820
|||||
QY 1145 GCGTTTGGAGCTTGGACAGATGCCAATTCACCTTTTGGATTTCTTAAACGCCACAGATG 1204
|||||
Db 1821 GCGTCTGAGCTTGGACAGATGCCAATTCACCTTTTGGATTTCTTAAACGCCACAGATG 1880
|||||
QY 1205 CCAGATGTGCTCCGCTCTGATCTCTTGTGAAGTGAATGAGGAGTGGCTTAATCTCTGG 1264
|||||
Db 1881 TCAGATGTGCTCCGCTCTGACACTTTGTTGAAGTGAATGAGGAGTGAATCTCTCTGT 1940
|||||
1265 G 1265
1941 G 1941

RESULT 6

US-10-425-114-6702
; Sequence 6702, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6702
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: 700575557_FLI
US-10-425-114-6702

Query Match 53.6%; Score 845.2; DB 13; Length 1200;
Best Local Similarity 94.8%; Pred. No. 5.9e-246;
Matches 874; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```

QY 450 TTCATATTCCTCCATGAACAGGGGATTCCTGTAGCTCTCTCTGGGACTATTGAGGT 509
DB 2 TTTTCATATTCCTCCATGAACAGGGGATTCCTGTAGCTCTCTCTGGGACTATTGAGGT 61
QY 510 CAATTTGTTGGTCTCTTAGCCCATTTGATTTTATCTATATTCCTGAGGCTTGAAGT 569
DB 62 CAATTTGTTGGTCTCTTAGCCCATTTGATTTTATCTATATTCCTGAGGCTTGAAGT 121
QY 570 CATGGCTCGAATCTTGACAGAGAGCAGCTTGAACACACACTATATCTGATGAAAGAG 629
DB 122 CATGGCTCGAATCTTGACAGAGAGCAGCTTGAACACACACTATATCTGATGAAAGAG 181
QY 630 GCTAAGCGGCAACTAATGAAGAAATGATGATGATGATGATGATGATGATGATGATG 689
DB 182 GCTAAGCGGCAACTAATGAAGAAATGATGATGATGATGATGATGATGATGATGATG 241
QY 690 CATGCCACCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749
DB 242 CATGCCACCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301
QY 750 TCTACAGTGCAGTATTTATTCATCATCATCATCATCATCATCATCATCATCATCAT 809
DB 302 TCTACAGTGCAGTATTTATTCATCATCATCATCATCATCATCATCATCATCATCAT 361
QY 810 CTTCATCATCTTCTGAAATTTGAAATGATGATGATGATGATGATGATGATGATGATG 869
DB 362 CTTCATCATCTTCTGAAATTTGAAATGATGATGATGATGATGATGATGATGATGATG 421
QY 870 AATTTGGCTATTTCTGAACCAACAGTGTCTTCAATCCGCTGGGTCTCGGGTTCGAAA 929
DB 422 AATTTGGCTATTTCTGAACCAACAGTGTCTTCAATCCGCTGGGTCTCGGGTTCGAAA 481
QY 930 ATTGGTGAATCTGAACAGTGTCTTCAATCCGCTGGGTCTCGGGTTCGAAA 989
DB 482 ATTGGTGAATCTGAACAGTGTCTTCAATCCGCTGGGTCTCGGGTTCGAAA 541
QY 990 GCGCTTAAGATGTTGTTCAAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1049
DB 542 GCGCTTAAGATGTTGTTCAAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 601
QY 1050 TCCCTGCTTGACACTTACTCTAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 1109
DB 602 TCCCTGCTTGACACTTACTCTAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 661
QY 1110 ACATATGTTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1169
DB 662 ACATATGTTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721
QY 1170 AATACACCTTTTGAATTTTAAAGCGGCGAGAGTGGAGTGGAGTGGAGTGGAGTGG 1229
DB 722 AATACACCTTTTGAATTTTAAAGCGGCGAGAGTGGAGTGGAGTGGAGTGGAGTGGAG 781
QY 1230 TTGCTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1289
DB 782 TTGCTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 841
QY 1290 GCTGGAGGCAAAAGTGTGGAGGGGATGATGATGATGATGATGATGATGATGATGATG 1349
DB 842 GCTGGAGGCAAAAGTGTGGAGGGGATGATGATGATGATGATGATGATGATGATGATG 901
QY 1350 AGCTTGTGAGGAGGAGGACA 1371
DB 902 AGCTTGTGATAGAGAGAGGA 923

```

RESULT 7

US-10-424-599-44858
Sequence 44858, Application US/10424599
Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 44858
LENGTH: 3077
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_140506C.1
US-10-424-599-44858

Query Match 35.9%; Score 565.2; DB 13; Length 3077;

Best Local Similarity 65.9%; Pred. No. 2.1e-160;
Matches 883; Conservative 0; Mismatches 448; Indels 9; Gaps 4;

```

QY 9 GAACATTTGCGGATGTTCTCTATCGAAGGCTGCCCACTGTATTTTCAGGCTATTTCAGT 68
DB 1019 GAACATTTGCGGATGTTCTCTATCGAAGGCTGCCCACTGTATTTTCAGGCTATTTCAGT 1078
QY 69 CTGCTCTCTGGGATTCACGAGTACAAATCTCTTCTGCAAGGAGTGGCGCATGATGAG 128
DB 1079 TTGCTACCCGGTACCACATCAGTACAGTCTTTTGTGATGAGAGTGGCGCATGATGAG 1138
QY 129 CGTCAACCTTACCATATCTGGGGAGTCTGGCATAGTAAACACACTTACTTTCACAAGGGA 188
DB 1139 CATCAACCTTATGCTGCTGGAGATATGGATAGTCTAACACTGT-CTTATTGGGCCAC 1197
QY 189 TATAACCAATTAACACCTTATCAAGTCCAGACACACTGGAAGCAGATGAAATGAT 248
DB 1198 TGATCTCTAAATACATACATCTGTTTACCTCCAG--ACGTGCTTCTGGAATAGCATGGAT 1255
QY 249 GTGATATGAATTTTCAACGTAACGTTTACGTTTGTGATGATGATGATGATGATGATG 308
DB 1256 GTGATATGATGCTTTTCCCGGATGCGCGGTTGACCGATGATGATGATGATGATGATG 1315
QY 309 ACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGTGTCTGTGTTCTGAAATATCTG 368
DB 1316 CTGCAAGATATCAGATACATCTGATGTACAAATATCCCGTCAGCGTATTTCTGCAATTCTA 1375
QY 369 AATTTGATACATGCTATGATTTACTCCAGATTTCTGGCAGGTTATTCCTAGACATT 428
DB 1376 TCTTCACACACCGCTTATGAAATTTACTCCCGAGTCAGGCAAGGTTGTGCTTTGGATGTT 1435
QY 429 AATTTACTCTGGAAGCAATCAATTCATATTCATGAAACAGGAGTTCCTGTAGCTCCT 488
DB 1436 GATCTACCACTGAAACAGGCAATTCATATTCATGAGCAGGAGTTCCTAGGCTCCT 1495
QY 489 CTCTGGGACTCATTCAGAGGTCATTTGTTGGTCTCCTTAGCCCATTTGGATTTTCATCTC 548
DB 1496 CTTTGGGACTCTCTGCAAGGGCAATTTGTTGGTCTCCTTAGTGTCTTGGATTTTATTTTA 1555
QY 549 ATATTCGGGAGCTAGAACTCATGCTCGCACTTGAACAAGCAGGCTTGAACAACAC 608
DB 1556 ATTTTAAGAGAGCTGGGGAATCATGGATCCAAATCTGACAGAAGAGGAGCTTGAACAACAT 1615
QY 609 ACTATATCTCATGGAAGAGGCTTAGCGGCAAACTAATGGAAGAAATGATGATGATGATG- 667
DB 1616 ACCATATCAGCTTGGAGAGGAGGAAATCTATCTAAATAGACAGACATGAGCATGGA 1675
QY 668 --GCGACCGCAACAGCATCTAGTGGATGCGACCCCTTATGATCTTTGAGGAGCATTTGCA 725

```

Db 1676 ACTGCATTTTCAAGATGTTTATCCATGCGAGGCGCATATGATATCTGAAAGATATTGGC 1735
 QY 726 GTAAAGCTTTTGCATAATGCAATTTCTACAGTGCAGTTATTTATTTATTCATCATCATCAGAT 785
 Db 1736 ATGAGATCTTTGCAAGAGAGGTTTCAACAGTTCCCTATATCCATTCATCTTCTGAAGAT 1795
 QY 786 GGAATCATCCCGAGTTATTTGCAATCTTTCGATCATCTTCTGGAATTTTGAATGATTTGT 845
 Db 1796 GCTTCATTTCCACAGTTACTACATCTTCTGCTTCACTTTTCAGGAATCTTAAATGCAATTTG 1855
 QY 846 AGATATCTTCAAAACTCAAGTGTAAATTCGCTATTTCTGAACCAACAGTGTCTCCATTT 905
 Db 1856 AGGTATTTTAGGCACGCTCTAGTTCTTCTGCTGACTTCAACTTCCAACTCTGCAATA 1915
 QY 906 CCCTCGGTTTCTGGGTTCCGAAATTTGGTGATCTGAACAGTGTCTCAATGGCTATTTG 965
 Db 1916 CTTGTGGCAGCTGGTGGCCCAAAATTTGGGGAATCAATCGCGGCTCTAGCAATGTTG 1975
 QY 966 CGACCTAATGCTTCACTTAGCTCTGCTTAAACATGTTGGTTCAAGCTGGAGTAAGCTCA 1025
 Db 1976 AGACCAACCGCTTCTCTTGGCTCAGCCCTAAATTTATTAGTTCAAGCCCAAGTAAGCTCA 2035
 QY 1026 ATACCAATTTGTCATGACAGACTCCCTGCTTGACACTTACTCTAGAAGTGACATCA 1085
 Db 2036 ATACCAATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2095
 QY 1086 GCCTTAGCAAAAGACAGAGTCTTACACACATGTTCCGCTGGATGAGATGACCATTTCCACAG 1145
 Db 2096 GCTTTGGCAAGACAGAGCATATACACATATTAATCTTGACCAATGACTGTTTCATCAG 2155
 QY 1146 GCTTTGACGTTGGCAAGATGCCATACACCTTTTGGATTTCTTAAGCGGCGAGATGC 1205
 Db 2156 GCTTTGACGTTGGCGGAGGAGCATATAGTCCCTATG---AGCTTAGAAGTCAAGATGT 2212
 QY 1206 CAGATGCTCGCTGGTCTGATCTTCTGCTGAAGTGTGAGAGGACTGGCTATCTCTGGG 1265
 Db 2213 CAGATGCTGGTCTGATCTTCTGCTGAAGTGTGAGAGGACTGGCTATCTCTGGG 2272
 QY 1266 GTGGCGGGGTGTTCAATTTGGAAGCTGGGAGCAAAAGTGTGGAGGGCATCATCACTA 1325
 Db 2273 GTGAGCGGCTGTGATCTGTTGAAGCTGGGAGCAAAAGTGTGGAGGGCATCATCACTA 2332
 QY 1326 AGTGACATTTTCAAGTCTT 1345
 Db 2333 AGTGACATTTCAAGTCTT 2352

RESULT 8

US-10-425-114-29527
 ; Sequence 29527, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 28
 ; CURRENT FILING DATE: 2003-04-28
 ; SEQ ID NO 29527
 ; LENGTH: 1507
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURES:
 ; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY057G02_FLI
 US-10-425-114-29527

Best Local Similarity 67.6%; Pred. No. 6.8e-143;
 Matches 743; Conservative 0; Mismatches 350; Indels 6; Gaps 2;
 QY 250 TGGATATGAAATTTTCAAGTACGTTAGTGTGATGAGTGGACCGTTTTCAGAGGTA 309
 Db 1 TGGATAATGATGCTTTTGGCGAATGCGCCGGTTGACCGATGGTACTTTGAGTGAAGTGC 60
 QY 310 CTCTGAGAGTTTTCAGAGGCTGCAATACAAATATCTAGTGTGCTGTTTCTGAATATCTGA 369
 Db 61 TGCCAAAGATATCAGATCTAGTGTACAAATATCCCGTCAAGCTATTTCTGCAATTTCTAT 120
 QY 370 ATTTGCAATGATGATGATTTTCTCCAGATTTCTGCAAGGTTTATGCGCTGACATTA 429
 Db 121 CTTTCAACACCGCTTATGAAATTTCTCCGAGTCAAGCAAGTGTGCTTGGATGTTG 180
 QY 430 ATTTTACCTGTGAACCAATCATTCATATTTCTCCATGAACAGGAGTTCCTCTAGCTCTTC 489
 Db 181 ATCTACAGTGAACAGGCTTTCTATATTTGCAAGTGAAGGAGTTTTCAGGCTCTCTC 240
 QY 490 TCTGGGACTCAATTCAGAGGTCATTTTGTGCTCTCTTAGCCCATTTGCAATTTCTACTCA 549
 Db 241 TTTGGGACTTCTGCAAGGCGCAATTTTGTGCTCTCTTAGGCTTCGATTTTATTTAA 300
 QY 550 TATTGCGGAGCTAGAACTCATGGCTCGAATTTGACAGAGAGCAGCTTGAACACACA 609
 Db 301 TTTTAAAGAGAGCTGGGGAATCATGGATCCAATCTGACAGAGAGGAGCTTGAACACATA 360
 QY 610 CTATATCTGATCGAAGAGGCTTAAGCGCAAACTAATGAGAAATGATAGTCAAGTG-- 667
 Db 361 CCAATCAGCTTGGAAAGAGGAAATCTGATCTTAAATAGACAGAACTTGAAGATGGAA 420
 QY 668 -GGAGCGCAACAGCTAGTGTGATGCTGCAACCCCTTATGATCTCTGAGGAGCATTTGCAG 726
 Db 421 CTGCAATTTTCAAGATGTTTATCCATGAGGSCCATATGATATCTGAAAGATTTGCCA 480
 QY 727 TAAAGCTTTTGAAGATGAGCTTTTCTACAGTGCAGTTATTTATTCATCATCATCAGATG 786
 Db 481 TGAAGATCTTGAAGAGGAGGTTTCAACTGCTTCTATTTATCCATTTCTTCTGAAGATG 540
 QY 787 GATCATTTCCGAGTTTATGCACTTTGCACTCTTCTGCAATTTTGAATGATTTTGA 846
 Db 541 CTTTCAATTTCCAGATTTACTACATCTTCTGCTTCACTTTTCAGGATTTCTTAATGCAATTTGTA 600
 QY 847 GATCATTTCAAAACTCAACTGTTAATTTGCTTCTTGAACCAACAGTGTGCTTCCATTC 906
 Db 601 GGTATTTTAGGCACCTGCTCTAGTGTCTGCTTCTGCTTCTCACTTCTCAATCTGTCATATC 660
 QY 907 CGCTGGGTTCTGGGTTCCGAAATTTGCTGATCTGAAACAGTGTCTCCATTTGGCTATGTTGC 966
 Db 661 CTGTGGGCAAGTGGTGGCGCAAAATTTGGGGAATCAAAATCGCGGCTCTTAGCAATGTTGA 720
 QY 967 GACCTAATGCTCCTCCTAGTCTGCTGCTTAACTGTTGGTTCAAGCTGGAGTAAAGTCAA 1026
 Db 721 GACCAACCGCTTCTTGTGCTGAGCTTAAATTTATTTAGTTTCAAGCCCAAGTAAAGTCAA 780
 QY 1027 TACCAATTTGGATGACAAAGCTCCCTGCTGACACTTCTCTAGAGTGCATCATCAG 1086
 Db 781 TACCAATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 QY 1087 CCTTAGCAAAAGACAGGCTTACACATGTTTGGTGGATGATGATGATGATGATGATGATGATGAT 1146
 Db 841 CTTTGGCAAAAGAACAGAGCATATACATATTTAATCTTGAAGAAATGATGATGATGATGATGATGAT 900
 QY 1147 CTTTGGAGCTTGGCAAGATGCAATATACACCTTTTGGATTTCTTAAAGCCAGAGATGCC 1206
 Db 901 CTTTGGAGTTGGCCAGGAGCATATAGTCCCTATG---AGCTTAGAGTCAAGATGTC 957
 QY 1207 AGATGCTCTCGGTTCTGATCTCTTGTCTGAAGGTTGATGAGGAGCTGCTTAATCTCTGGGG 1266
 Db 958 AGATGTTTGGGTTCTGATCTCTTGTCTGAATAAAGTGTGAAAGCTTGGCAAAATCCAGGTTG 1017
 QY 1267 TCGCGGCGGTTCTCTTGTGGAAGCTGGGAGCAAAAGCTGAGGAGCATCATATCACTAA 1326

Query Match 32.2%; Score 507; DB 13; Length 1507;

Db 1018 TCAGGGGGCTTGTCATCTGCGAAGCTGGCAGCAGCGGTGAGAGGCAATGTTTCATTGA 1077

Qy 1327 GTGACATTTTCAAGTTCTT 1345

Db 1078 GTGACATATTCAGATTCTT 1096

RESULT 9

US-10-424-599-79689
 ; Sequence 79689, Application US/10424599
 ; Publication No. US2004031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 79689
 ; LENGTH: 2315
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_42975C.1
 US-10-424-599-79689

Query Match 32.0%; Score 504.4; DB 13; Length 2315;
 Best Local Similarity 62.4%; Pred. No. 5.6e-142; Indels 48; Gaps 2;
 Matches 852; Conservative 0; Mismatches 466;

Qy 8 GGAACATTTCCGATGCTCTCTATCGAAGGCTGCCCACTGTAATTCAGGCTAATTCAG 67

Db 581 GAACCATATAGCTATGCTCCGATGGAGGATGCCCTGCTGTGTTTCAAGTTGTTGCAA 640

Qy 68 TCTGCTCTCGGATTCACGAGTACAAATTTTGTGGACGGGAATGGGGCATGATGA 127

Db 641 CTTAATGCCGGGTTTCATCAGTACAAATTTAATGTAGATGGTGGTGGGGCAGATGA 700

Qy 128 GCGTCAACCTACCATATCTGGGGAGTTGGCATAGTTTAAACACTTTTACTTGAACAAGGA 187

Db 701 SCAGACCCATTTGTAATGGAGCTGTGGAGTAGTGAACACTATTTATATAGTGAGA 760

Qy 188 ATATAACAAATAAACACCTTATCAAGTCCAGACACCTCGGAAGCAGGATGAACATGGA 247

Db 761 ACCAGATATCTTACCTTCCATCTTAATATCTGAAACACCTGGTCGATCACACATGGAGGT 820

Qy 248 TGTGATATGMAAATTTTCAACGTACGGTTACGTTGTGATGGCAGCCGTTTCAGAGG 307

Db 821 TGACACATG-----GAGC 835

Qy 308 TACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGCTGTGTTTCTGAATATCT 367

Db 836 TAATCCAGGATGACTGTGCTGATCTGGAGTTCTGCCACCGTATATCGTATTCCT 895

Qy 368 GAATTTGATACATGCTATGATTTACTCCAGATTTCTGGCAAGTTATTCGCTAGACAT 427

Db 896 GTCCACGCTACTTCAATATGATTTGCTCCGAGTCAGGAAAGGTCAITTCGCTGGATAT 955

Qy 428 TAATTTACTCTGAGGATATCATTCATATCTCCATCAAGCAGGGATTCCTGTAGCTCC 487

Db 956 AAATTTACCACTTAAGCAAGCATTCATGTTCTTTTACCAAGGATATCTATGCGCTCC 1015

Qy 488 TCTCGGACTCATTCAGAGGTCAATTTGTTGGTCTCCCTAGCCCAITGGATTTTCATCT 547

Db 1016 TCTATGGATTTTTCAGAGCCAGCTGTTGGAGTTCTTAGTCAATGACTTCAITCT 1075

Qy 548 CATATTGGGAGCTAGAACTCATGCTCGAACTTGTGACAGAGAGCGCTTGAACA 607

Db 1076 AATATTGAAGAGCTGGGGAACCAAGTTCAAATTTGACTCAAGAACTCGAGACTCA 1135

Qy 608 CACTATATCTGCATGGAAGAGGCTAGCGGCAAACTAATGGAAGAAATCATAGTCAGTG 667

Db 1136 TACTATAGCTGCTCGAAGAGAGAAATTTCAACATTCAGAACACTTGAATAGTATGG 1195

Qy 668 GCGACCGCAAC---AGCATCTAGTGCATGCCACCCCTTATGAGTCCCTGAGGACATGCG 724

Db 1196 GGGATCATATCTCTGGCGATTTGTTTCATGCCGACCCCATGAATGTCTAAAGATGGC 1255

Qy 725 AGTAAAGCTTTTGCAGAAATGCGCATTTCTACAGTGCAGCTTATTTATCATCATCAGAGA 784

Db 1256 TTTGAAGGTTTTCAGAAACAAAGGTGCAACTGTCCTCATCATCATCTTCTTCAGAGGA 1315

Qy 785 TGGATCATTTCCCGCAGTTATTGTCATCTTGCATCACTTCTTCGGAATTTTGAATGATTTG 844

Db 1316 TGGTTTATATCTCTCACTGCTCATCTTGTCTCCCTATCCGGAATCTATAAGGATATG 1375

Qy 845 TAGATACTTCAAAAACCTCAACTGGTAAATTTGCTTATTTCTGAAACCAACAGTGTCTCAT 904

Db 1376 CAGGCATTTTAAAGCACTCTCTGAGTTCTTTGCCCATTTCTCACTCCAGTGGCTTCAAT 1435

Qy 905 TCCGCTGGGTTCTCGGTTCCGAAATTTGGTATCTGAACAGTCTCCATTTGGCTATGTT 964

Db 1436 ACCTTTGGGTACATGGATGCTTAGAGTTGGGGAACCAAAATGGTGGGCCACTAGCAATGTT 1495

Qy 965 GCGACCTAATGCTCACTTAGCTCTGCCCTTAAACATGTTGGTTCAAGCTGGAGTAAGCTC 1024

Db 1496 GATGCCAAGTGTCTCTCGTGTCTTGTCTATGTTGTTCAAGCCGAGTTAGCTC 1555

Qy 1025 AATAACAAATTTGGATGACAAACGCTCCCTGCTGTGACACTTACTTGAAGTGCATGAC 1084

Db 1556 AATACCGATTTGGATGCTAATGATTCGTTGCTTGACATTTATTCAGAGAGCGATATATAC 1615

Qy 1085 AGCCTAGCAAGAGCAAGGCTTACACACATGTTGGCTGGATGATGATGACCATTCACCA 1144

Db 1616 TGCATTTGGCTTAAAGATAAAGCTTATGCTAGATATCTCTCGATGAATTTAGTATTCACCA 1675

Qy 1145 GGCTTTGCGAGCTTGGCAAGATGCCAATAACACCTTTTGGATTTCTTAAACGCCAGAGATG 1204

Db 1676 GGCATTTGCTTTTGGCAAGATGCACTTCTCTTATGCTGCTTCAATGCTCACAGATG 1735

Qy 1205 CCAGATGTGCTCCGCTGTGATCTCTGCTGAAAGGATGAGGAGCTGGCTTAATCTCTGG 1264

Db 1736 TCACATGTTTGGATCTGATTCGCTTTTCAAAAGTCAATGAGCGGTTGGCTAATCTCTGG 1795

Qy 1265 GGTGCGCGGGTGTTCATTTGTTGAGCTGGAGCAAAAGTGTGAGGCGCATCATATCACT 1324

Db 1796 GGTAGGAGACTTGTGGCTGTGGAGGCTGGCAGCAAGCGGTGGAAGGGATTTTCTTT 1855

Qy 1325 AAGTGACATTTTCAAGTTTCTTCTGCTGAGCTTGTGAGGAGAGGAGC 1370

Db 1856 AAGCGATCTGTTTCAGATTTCTTTAGGCTAGTTGGGATATATGGC 1901

RESULT 10

US-10-437-963-90730
 ; Sequence 90730, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966

Query Match 16.9%; Score 266.2; DB 9; Length 282;
Best Local Similarity 98.6%; Pred. No. 4.7e-70;
Matches 279; Conservative 0; Mismatches 3; Indels 1; Gaps 2;
QY 86 CGAGTACAAATCTTTGTGGACGGGAATGGCGGATGATGAGCGTCAACCTTACCATATC 145
DB 1 CGAGTACAAATCTTTGTGGACGGGAATGGCGGATGATGAGCGTCAACCTTACCATATC 60
QY 146 TGGGAGTTTGGCATAGTTAAACACATTTACTTGCACAGGGAATATPAACCAATTAACAC 205
DB 61 TGGGAGTTTGGCATAGTTAAACACATTTACTTGCACAGGGAATATPAACCAATTAACAC 120
QY 206 CTTATCAAGTCCCAAGCACACCTGGAGAGGAGGATGATGAGTGGATGATGAGTAAATTT 265
DB 121 CTTATCAAGTCCCAAGCACACCTGGAGAGGAGGATGATGAGTGGATGATGAGTAAATTT 180
QY 266 TCAACGTACGTTACGTTCTGATGAGTCCGCTTTCAGAGGATCTGAGAGTTTCAGA 325
DB 181 TCAACG-GGGGTACGTTCTGATGAGTCCGCTTTCAGAGGATCTGAGAGTTTCAGA 239
QY 326 GGCTGCAATACAAATATCTAGTGTGCTGTTTCTGAATATCTG 368
DB 240 GGCTGCAATACAAATATCTAGTGTGCTGTTTCTGAATATCTG 282
RESULT 15
US-10-437-963-66100
; Sequence 66100, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbarov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 66100
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_67084C.1
US-10-437-963-66100
Query Match 15.3%; Score 241.6; DB 17; Length 1402;
Best Local Similarity 54.0%; Pred. No. 4.3e-62;
Matches 552; Conservative 0; Mismatches 449; Indels 21; Gaps 2;
QY 321 TCAGAGCTGCAATACAAATATCTAGTGTGCTGTTTCTGAATATCTGAATTCGATACA 380
DB 343 TCAGAGCTGCAATACAAATATCTAGTGTGCTGTTTCTGAATATCTGAATTCGATACA 402
QY 381 TGCTATATATCTCCAGATCTGGAGAGGATTTATCCCTAGACATTAATTTACCTGTG 440
DB 403 ATATATGATGTGTTCTGTTTCTAGCAAGATAGCAGTTTGGATGCTGCTCTCTCTGTC 462
QY 441 AAGCAATCATTCATATCTCCATGACAGAGGATTCCTGTAGTCTCTCTCTGGACTCA 500
DB 463 AAGCAAGCTTTAAATCATGATGATGAGGATCTTTCTTGGTTCCTCTTGGATGAT 522
QY 501 TTCAGAGTCAATTTGTTGTTCTCTAGGAGGATTTGATTTCACTACTCATATTTGGGAG 560
DB 523 CAGCAGCAAAACCGTAAACAGGATGCTAAACCGATCCGATTTGTTGTTTAAATTTGAGAAAG 582
QY 561 CTAGAACTCATGCTCGAATCTGACAGAGAGGAGCTTTGAACACACACATATATCTGCA 620

DB 583 CTGACAGAAACATTCGAATCTTGGCCATGAGGAATTTGAATGATTCAGTGTCTGCT 642
QY 621 TGGAAAGAGCTTAAGCGCAAACTAATGGAAGAAATGATAGTCACTGAGGCGACCGACAG 680
DB 643 TGGAAAGAGCAAAATTCAGATTTTATGGGGGACCTGATGTTGCTGCCATTCAGAAAGG 702
QY 681 CATCTAGTGCATGCCACCTTATGATGCTTGGAGGACATTCAGTAAAGCTTTTGCAA 740
DB 703 CCATTAATTCATTTAAGGATTCAGATAATTTAAGGATGTCGATTTGGCTATATATAGA 762
QY 741 AATGGCAATTTACAGTGCAGTATTTATTCATCATCATCATCATCATCATCATTCOCGAG 800
DB 763 AATGAATATCTTCAGTCC--TATCTTTAAGCCCTCAAGGATTCATCAGGATGCCT 819
QY 801 TTATTCATCTTTCGATCATCTTCTGGAATTTTGAATGATTTTGTAGATATCTTCAAAAC 860
DB 820 TTGCTTGGTCTTGCACACCTTCCAGGATTTGAAATTTATTTGCTCAAAAGCTGCAAGAA 879
QY 861 TCAACTGTAATTTGCTTATCTGAAACCAACAGTGTGCTCCATTCGGCTGGGTTCTGCG 920
DB 880 CAGCCTGAAAGGATCTCATTTCTGCAAAACAGATTCAGTATGCTTATGTTACATGG 939
QY 921 GTTCCGAAATTTGCTGATCTGAAACAGTGTGCTCCATTTGCTGATTTGGGACCTTAATG 980
DB 940 TCACCATATCTGCGAAGCAAGCAATAGACAGCTTAGAATCTTCGCGACCAAGCACTCCT 999
QY 981 CTTAGCTTGCCTTTAATCATGTTGTTCAAGCTGAGTGAATCAATACCAATTTGTGAT 1040
DB 1000 CTAAATTCATGCTGATTTGCTGTTTGAAGATAGATAGTCAATTTCTATATAGTTGAC 1059
QY 1041 GACACGACTCTGCTGTTGACACTTACTCTAGAGTGACATTCACAGCCTAGCAAAAGAC 1100
DB 1060 GATAAAGGCGCTCTCTCTGATGCTTACTCGCTGATGATTCATGCTTATGCTTAGGCA 1119
QY 1101 AAGTCTTACACACATGTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1160
DB 1120 GATGCTTACACTCGTATTGAGCTTGAACAGTGAACGCTGAGCATGCTTGGAGCTGCAA 1179
QY 1161 CAAGATCCCAATACACCTTTTGGATTTCTTTAAGCGGACAGATGCGAGATGCTGCTCGG 1220
DB 1180 TACCAGGTGAA-----TGGCGGAAAGACACTGTTCATACCTGCTTGAGC 1221
QY 1221 TCTGATCCTTTGCTGAAGGTGATGAGCGATGCTGCTTAATCCTGCTGCTGCTGCTGCTG 1280
DB 1222 ACTAGTACCTTCTTGGAGTTTGGAGCAATTTGTCAGTCCAGGGTGGCGGAGTCTC 1281
QY 1281 ATTGTGAAGCTGGGAGCAACAGTGTGAGGCGATCATATCATTAAGTGAATTTTCAG 1340
DB 1282 GTTATTGAACCAAGGAGCAGATTTGTGCAAGGAATTAATCTCATTTGAGGAGCGCATTTACA 1341
QY 1341 TT 1342
DB 1342 TT 1343

Search completed: July 9, 2004, 08:27:50
Job time : 783 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:46:38 ; Search time 59 Seconds
(without alignments)
2159.812 Million cell updates/sec

Title: US-09-857-525C-2
Perfect score: 2342
Sequence: 1 TREHLPMSPIEGPTVFQA.....SKRVEGIISLDIFKFLLSL 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	451	3 AAY96783	Partial Z
2	2085	89.0	493	3 AAY96784	Aay96784 Rice suc
3	2019	86.2	442	3 AAY96789	Aay96789 Wheat suc
4	1581	67.5	492	3 AAY96787	Aay96787 Soybean s
5	1525.5	65.1	482	3 AAY96786	Aay96786 Soybean s
6	1214	51.8	368	5 ABB90919	Abb90919 Herbicida
7	392	16.7	189	3 AAY96785	Aay96785 Parital r
8	373	15.9	328	2 AAW88438	Aaw88438 Disease a
9	373	15.9	328	4 AAB93432	Aab93432 Human ins
10	373	15.9	328	5 AAO18496	Aao18496 Human ins
11	373	15.9	352	5 AAO18495	Aao18495 Human ins
12	369	15.8	77	3 AAY96793	Aay96793 Wheat suc
13	367	15.7	331	2 AAW29817	Aaw29817 Mammalian
14	367	15.7	331	5 AAO18497	Aao18497 Human ins
15	367	15.7	331	6 ABU04258	Abu04258 Human exp
16	367	15.7	331	6 ABU04261	Abu04261 Human exp
17	367	15.7	331	6 ABU04262	Abu04262 Human exp
18	367	15.7	331	6 ABU04257	Abu04257 Human exp
19	367	15.7	331	7 ADE62623	Ade62623 Human pro
20	367	15.7	344	3 ABE54009	Abe54009 Human pan
21	367	15.7	344	6 ABU04259	Abu04259 Human: exp
22	361.5	15.4	634	4 ABB66245	Abb66245 Drosophil
23	361.5	15.4	634	5 ABG70033	Abg70033 Larval vi
24	355.5	15.2	330	7 ADE62621	Ade62621 Rat Prote
25	340.5	14.5	305	4 AAE00329	Aae00329 Human Prk

26	340.5	14.5	305	4 AAE00328	Aae00328 Human Prk
27	340	14.5	464	4 AAE00223	Aae00223 Human AMP
28	340	14.5	489	4 AAB47679	Aab47679 PRKAG3. 1
29	340	14.5	489	5 ABP43929	Abp43929 AMP activ
30	340	14.5	489	6 AAE32034	Aae32034 Human Kin
31	339.5	14.5	305	4 AAE00221	Aae00221 Human AMP
32	335.5	14.3	305	4 AAE00226	Aae00226 Sus scrof
33	335.5	14.3	305	4 AAE00225	Aae00225 Sus scrof
34	335.5	14.3	333	5 ABB06101	Abb06101 Human NS
35	335.5	14.3	464	5 AAE22987	Aae22987 Pig PRKAG
36	335.5	14.3	464	5 AAE22988	Aae22988 Pig PRKAG
37	334.5	14.3	305	4 AAE00220	Aae00220 Pig AMPK
38	334.5	14.3	464	4 AAE00222	Aae00222 Pig wild-
39	334.5	14.3	464	5 AAE22984	Aae22984 Pig wild-
40	334.5	14.3	464	5 AAE22985	Aae22985 Pig PRKAG
41	334.5	14.3	464	5 AAE22986	Aae22986 Pig PRKAG
42	334.5	14.3	514	4 AAE00224	Aae00224 Sus scrof
43	329	14.0	489	7 AAE38410	Aae38410 Human Prk
44	303.5	13.0	322	6 ABR52942	Abr52942 Protein s
45	302	12.9	488	4 ABG20078	Abg20078 Novel hum

ALIGNMENTS

RESULT 1

AAY96783

ID AAY96783 standard; protein; 451 AA.

AC AAY96783;

DT 26-SEP-2000 (first entry)

DE Partial Z. mays sucrose non-fermenting 4 protein.

KW Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
KW carbon catabolite repression; development; nitrogen partitioning.

XX Zea mays.

XX WO200036116-A2.

XX 22-JUN-2000.

XX 15-DEC-1999; 99WO-US029825.

XX 16-DEC-1998; 98US-0112564P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Allen SM, Heppard EP, Miao G, Weng Z;

XX WPI; 2000-431594/37.

XX N-PSDB; AAA51347.

XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
involved in carbon catabolite repression in plants and seeds, useful for
controlling carbon and nitrogen partitioning pathways during plant growth
and development.

XX Claim 10; Page 31-33; 48pp; English.

XX This is a partial maize (Zea mays) sucrose non-fermenting 4 (SNF4)
protein which is involved in carbon catabolite repression in plants and
seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
Arabidopsis thaliana and Saccharomyces cerevisiae, from clone
csk1c.pk001.c15 prepared from corn unpollinated developing silk 24 hours
after emergence. The polynucleotides are used in plants to control carbon
and nitrogen partitioning pathways during plant growth and development.
XX The catabolite repression proteins would facilitate studies for better
understanding the mechanism of catabolite repression in plants and could
provide genetic tools to enhance or otherwise alter the accumulation of
carbohydrates, lipids and proteins during plant growth and development

XX PF 15-DEC-1999; 99WO-US029825.
 XX PR 16-DEC-1998; 98US-0112564P.
 XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PI Allen SM, Heppard EP, Miao G, Weng Z;
 XX DR MPI; 2000-431594/37.
 XX DR N-PSDB; AAA51353.
 XX XX
 PT New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 XX ES Claim 24; Page 44-45; 48pp; English.
 XX XX
 CC This protein is a wheat sucrose non-fermenting 4 (SNF4) protein which is
 CC involved in carbon catabolite repression in plants and seeds. The cDNA
 CC was isolated, based on similarity to SNF4 proteins from Arabidopsis
 CC thaliana and Saccharomyces cerevisiae, from library wll prepared from
 CC wheat leaf from 7 day old seedling. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development
 XX SQ Sequence 442 AA;

Query Match 86.2%; Score 2019; DB 3; Length 442;
 Best Local Similarity 86.5%; Pred. No. 3.3e-196;
 Matches 384; Conservative 33; Mismatches 25; Indels 2; Gaps 1;
 QY 7 MSPIEGCTVFOAICSLSPGHEHYKFFVGDGWRHDERQPTISGEFGVNTLYLTREYNQI 66
 DB 1 MSPVEGCTVFOAICNLPPIGYKFNVDGQWRHDEGQPTITGEYGVNTLYLTREFDHI 60
 QY 67 NTLSSPSTPGSRMMDVDNENFQRTVTLSPGTSEGLRVSEAAIQISCRVSEYLNHT 126
 DB 61 NTVLSTPTFSR--MDVDSDFSQMSLSDGALQEGSPRISEAAIQISCRVAEYLNHT 118
 QY 127 CYDLLPDSGKVIALDINLPVKQSFHILHEQGPVAPLWDSFRGQFVGLLSPILDFILIRE 186
 DB 119 GYDLLPDSGKVIALDINLPVKQSFHILHEQGPVAPLWDSFRGQFVGLLSPILDFILIRE 178
 QY 187 LETHGSLNLTETHELTISAWKEAKQTNGRNDSQWRPOHLVHATPYESLRDIAVKLLQ 246
 DB 179 LETHGSLNLTETHELTISAWKEAKQTYGRNDGQLRSNQHLVHATPYESLRGIAKMLE 238
 QY 247 NGISTVPYIYSSSDGSPFQLLHLASLSGILKICRYFKNSTGNLPILNQPCVSLPGSW 306
 DB 239 TGISTVPYIYSSSDGSPFQLLHLASLSGILKICRYFKNSTGNLPILNQPCVSLPGTW 298
 QY 307 VPKIGDLSNRPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSDLDITYSRSDITALAKD 366
 DB 299 VPKIGEPNGHPLAMLRPNLTSSALNMLVQAGVSSIPVDDNDSDLDITYSRSDITALAKD 358
 QY 367 KYTVHRLDEMTIHOALQIQDANTPFQFNQRCOMCLRSDFLKWMLRANPGVRRVF 426
 DB 359 KYTVHRLDEMTIHOALQIQDANSFGLFNQRCOMCLQSDFLKXWERLANPGVRRVF 418
 QY 427 IVEAGSKRVEGIISLIDFKLLS 450
 DB 419 IVEAGSKRVEGISLIDFKLLS 442
 RESULT 4
 AAY96787
 ID AAY96787 standard; protein; 492 AA.

XX AC AAY96787;
 XX DT 26-SEP-2000 (first entry)
 XX DE Soybean sucrose non-fermenting 4 protein from clone sf11.pk0004.b.
 XX KW Soybean; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 XX KW carbon catabolite repression; development; nitrogen partitioning.
 XX OS Glycine max.
 XX XX
 XX PN WO2000036116-A2.
 XX PD 22-JUN-2000.
 XX PF 15-DEC-1999; 99WO-US029825.
 XX PR 16-DEC-1998; 98US-0112564P.
 XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PI Allen SM, Heppard EP, Miao G, Weng Z;
 XX DR MPI; 2000-431594/37.
 XX DR N-PSDB; AAA51351.
 XX XX
 PT New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 XX PS Claim 10; Page 40-41; 48pp; English.
 XX XX
 CC This is a soybean (Glycine max) sucrose non-fermenting 4 (SNF4) protein
 CC which is involved in carbon catabolite repression in plants and seeds.
 CC The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae, from library sfl
 CC prepared from soybean immature flower. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development
 XX SQ Sequence 492 AA;

Query Match 67.5%; Score 1581; DB 3; Length 492;
 Best Local Similarity 67.6%; Pred. No. 1.6e-151;
 Matches 305; Conservative 61; Mismatches 75; Indels 10; Gaps 4;
 QY 3 EHLMPSPGCTVFOAICSLSPGHEHYKFFVGDGWRHDERQPTISGEFGVNTLYLTRE 62
 DB 46 ELLMPSPGCTVFOAICNLPPIGYKFNVDGQWRHDEGQPTITGEYGVNTLYLTRE 105
 QY 63 YNOITLSSPSTPGSRMMDVDNENFQRTVTLSDGTSEGLRVSEAAIQISCRVSEYL 122
 DB 106 PNYMPVL--PPDVASGNSMDVDNDAFRMARLTDGTLSEVLPRISDQVTSRQISAF 163
 QY 123 NLHTCYDLLPDSGKVIALDINLPVKQSFHILHEQGPVAPLWDSFRGQFVGLLSPILDF 182
 DB 164 SSHTAYELLPEGSKVVALDVLFPVQAFLHLEQGVFMALPWFDFCKGQFVGLSASEFIL 223
 QY 183 ILRELATHGSLNLTETHELTISAWKEAK----RQTNGRNDSQWRPOHLVHATPYESLR 238
 DB 224 ILRELATHGSLNLTETHELTISAWKEGKSYLNQNGHGTAFSR---CFIHAGFYDNLK 280
 QY 239 DIAVKLLQNGISTVPYIYSSSDGSPFQLLHLASLSGILKICRYFKNSTGNLPILNQPV 298
 DB 281 DIAVKLLQNGISTVPYIYSSSDGSPFQLLHLASLSGILKICRYFKNSTGNLPILNQPV 340
 QY 299 CSPIGSWVPKIGDLSNRPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSDLDITYSR 358

```

Db 341 CAIPVGVTPKIGSNRPPLAMLRRTASLASAKLLVQAVSSIPVDDNDLLDIYCRS 400
QY 359 DITALAKDKVYTHVRLDEMTIHQALQLGQDANTPFQFFNGRCQWCLRSDDPLLKWERLA 418
Db 401 DITALAKNRAYTHINLDMTIVHQAALQGGDAYSPE-LRSQCQXCLRSDDPLLKWERLA 459
QY 419 NPGVRRVIVEAGSKRVEGIISLSIDIFKELL 449
Db 460 NPGVRLVIVEAGSKRVEGIISLSIDIFKELL 490

RESULT 5
AY96786
ID AY96786 standard; protein; 482 AA.
XX
AC AY96786;
XX
DT 26-SEP-2000 (first entry)
XX
DE Soybean sucrose non-fermenting 4 protein from clone ses4d.pk0040.b1.
XX
KW Soybean; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
KW carbon catabolite repression; development; nitrogen partitioning.
XX
OS Glycine max.
XX
PN WO200036116-A2.
XX
PD 22-JUN-2000.
XX
PF 15-DEC-1999; 99WO-US029825.
XX
PR 16-DEC-1998; 98US-0112544P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Heppard EP, Miao G, Weng Z;
XX
WPI; 2000-431594/37.
DR N-PSDB; AAA51350.
XX
New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
PT involved in carbon catabolite repression in plants and seeds, useful for
PT controlling carbon and nitrogen partitioning pathways during plant growth
PT and development.
XX
PS Claim 10; Page 37-39; 48pp; English.
XX
This is a soybean (Glycine max) sucrose non-fermenting 4 (SNF4) protein
CC which is involved in carbon catabolite repression in plants and seeds.
CC The cDNA was isolated, based on similarity to SNF4 proteins from
CC Arabidopsis thaliana and Saccharomyces cerevisiae, from library ses4d
CC prepared from soybean embryogenic suspension 4 days after subculture. The
CC polynucleotides are used in plants to control carbon and nitrogen
CC partitioning pathways during plant growth and development. The catabolite
CC repression proteins would facilitate studies for better understanding the
CC mechanism of catabolite repression in plants and could provide genetic
CC tools to enhance or otherwise alter the accumulation of carbohydrates,
CC lipids and proteins during plant growth and development
XX
SQ Sequence 482 AA;

Query Match 65.1%; Score 1525.5; DB 3; Length 482;
Best Local Similarity 67.5%; Pred. No. 7.1e-146;
Matches 303; Conservative 57; Mismatches 72; Indels 17; Gaps 6;

QY 3 EHLPMSPICPTVFOAICSLSPGIEHYKFFVGVGWRHDERQPTISGEGIVNTLYLIRE 62
Db 45 ELLQMSFVGGCTVFOVHSLVFGHHQYKFFVGVGWRHDDLQPCSEGEIVNTVSLATD 104
QY 63 YNQINTLSPSPFGSNMMDVNMENFQRTVTLSDGTVSEGL-RVSEAAIQISRCRVSEY 121

```

```

Db 105 PNILPVLTPEDIVSGS--NMDVDNEAFRRMVRLLTDTLSNVLLPRISDVIDTQTSQRISAF 162
QY 122 LNLHTCYDLLPDSGKVIALDINLPVKQSPHILHEQGIPIVAPLWDSFRGQFVGLSLPLDFI 181
Db 163 LMSSTAYELLPSGKVVLDVLPVQAAPHILHEQGIPIAPLWDICKGQFVGVLSDLDFI 222
QY 182 LILRELTGNSLTBEQLETHYTSIAWKEAKROTNGRNDSONRP-QQHLVHATPYESLRDI 240
Db 223 LILRELGNHGSNLTELETHYTSIAWKGK-----WGFTQCF-RAGPYDNLKEI 272
QY 241 AVKLLONGISTVPVIYSSSDGSPQLLHLASLSGLKICICRYFRNSTGNLPIKQPVCS 300
Db 273 AVKILQHGISTVPIIH--SEDSGFPQLLHLASLSGLKICICRYFRNCSSSLPIQLPICA 330
QY 301 IPLGSWVPKIGDLNSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNDLLDITYSRDI 360
Db 331 IPVGTWVPKIGESNRPPLAMLRPNASLSALNMLVQAGVSSIPVDDNDLLDITYSRDI 390
QY 361 TALAKDKVYTHVRLDEMTIHQALQGGDANTPFQFFNGRCQWCLRSDDPLLKWERLAMP 420
Db 391 TALAKDRTYTHINLDMTIVHQAALQGGDSYNTYB-LSCQRCQWCLRTDLSLHKVMERLASE 449
QY 421 GVRRLVIVEAGSKRVEGIISLSIDIFKELL 449
Db 450 GVRRLVIVEAGSKRVEGIISLSIDIFKELL 478

RESULT 6
ABB90919
ID ABB90919 standard; protein; 368 AA.
XX
AC ABB90919;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 130.
XX
KW Herbicidally active polypeptide SEQ ID NO 130.
XX
OS Herbicidal; plant; agriculture; herbicide.
XX
PN Arabidopsis thaliana.
XX
WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
(PARB) BAYER AG.
XX
Tietjen K, Weidler M;
XX
WPI; 2002-269010/31.
XX
Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
C:aim 5; SEQ ID NO 130; 261pp + Sequence Listing; English.
XX
The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
XX herbicides
XX
SQ Sequence 368 AA;

```

Query Match	16.7%	Score 392;	DB 3;	Length 189;
Best Local Similarity	44.2%;	Pred. No. 4.6e-31;		
Matches	84;	Conservative 34;	Mismatches 64;	Indels 8; Gaps 2;
QY	261	DGSEFPOLLHLAGSGILKICIFCFKNSGTGNLPILINQPVCSIPLGSGVKKPIGDLNSRPLAM	320	
DB	6	DSSGMPLLGLATLPGLVKTCSKLQCPQCYGFLQNGIVSMPTGTNSPHTGKASNRQURT	65	
QY	321	LRPNASLSGALNMLVQAGVSSIPFVDNDSLLPTYSRSDITALAKDKVYTFVRLDMMTH	380	
DB	66	SRPSTPLNSCLDILLEDVRSSIPFVDNGALLDVSLSIDIMALGKNDVYTRIELEQVTVE	125	
QY	381	QALQLGODANTPGFENQGR-CQMCILRSDPLLKVMERLANPGYRVRVIVEAGSRFVGII	439	
DB	126	HALELOQYV-----NGRRHCHTCLSTSTFLEVLQLSAPGVRVWVVISFRSFRVGGII	178	
QY	440	SLSDIFKFLI	449	
DB	179	SLRDATFELI	188	

RESULT 8
AAW88438
ID AAW88438 standard; protein; 328 AA.

XX	AAW88438;
XX	AC
XX	26-APR-1999 (first entry)
XX	Disease associated protein kinase DAPK-7.
DE	

DAPK-7: disease associated protein kinase; human; diagnosis; therapy; adult respiratory distress syndrome; allergy; asthma; arteriosclerosis; bronchitis; emphysema; hyperosinophilia; myocardial inflammation; pericardial inflammation; anaemia; rheumatoid arthritis; Addison's disease; AIDS; atherosclerosis; atopic dermatitis; dermatomyositis; diabetes mellitus; glomerulonephritis; gout; Grave's disease; lupus erythematosus; multiple sclerosis; myasthenia gravis; osteoarthritis; osteoporosis; pancreatitis; polycystic kidney disease; polymyositis; scleroderma; Sjögren's syndrome; autoimmune thyroiditis; cancer; infection; trauma; cell proliferation.

PR 16-DEC-1998; 98US-0112564P.

KW cell proliferation.

SO Homo sapiens.

FH	Key	Location/Qualifiers
FT	Modified-site	72

FT	
FT	/note= "potential cAMP dependent protein kinase site"
FT	77..91
Peptide	
FT	/note= "observed in many other organisms; also in "
FT	

Peptide	217..224	/note="conserved in AMPK gamma isoform and DAPK-7"
FT	217	Modified-site
FT	220..228	/note="potential cAMP dependent protein kinase site"

PS Claim 10; Page 36; 48pp; English.

FT
XX
PN WO9858052-A2.

XX
PD 23-DEC-1998.
XX
PE 19-JUN-1998:

19-JUN-1998; 98WO-US012813.


```
PR 12-JAN-2001; 2001US-0261457P.
PR 12-JAN-2001; 2001US-0261458P.
PR 12-JAN-2001; 2001US-0261459P.
PR 12-JAN-2001; 2001US-0261461P.
PR 12-JAN-2001; 2001US-0261518P.
PR 12-JAN-2001; 2001US-0261518P.
PR 12-JAN-2001; 2001US-0261531P.
PR 12-JAN-2001; 2001US-0261532P.
PR 12-JAN-2001; 2001US-0261589P.
PR 12-JAN-2001; 2001US-0261590P.
PR 12-JAN-2001; 2001US-0261634P.
PR 12-JAN-2001; 2001US-0261635P.
PR 12-JAN-2001; 2001US-0261697P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Seidel-Dugan C, Ferguson KC, Kidd T;
XX
XX WPI; 2002-599664/64.
XX N-PSDB; AAL48614.
XX
XX Identifying an insulin receptor signaling modulator, useful as drug
XX targets for treating diabetes or metabolic disorders, comprises
XX contacting an assay system comprising insulin receptor signaling
XX modifiers with a test agent.
XX
XX Disclosure; Page 49-50; 232pp; English.
XX
XX The present invention relates to a method of identifying a candidate
XX insulin receptor (INR) signaling modulating agent, involving contacting
XX an assay system comprising an insulin receptor signaling modifier (ISM)
XX polypeptide or nucleic acid with a test agent, and detecting a test agent
XX biased activity of the assay system. The method is useful for
XX identifying candidate INR signaling modulating agents. ISM genes may be
XX used as drug targets for treatment of disorders related to INR signaling
XX such as diabetes or metabolic syndrome. ISM nucleic acids and
XX polypeptides are useful for identifying and testing agents that modulate
XX ISM function and for other applications related to the involvement of ISM
XX in INR signaling, and for identifying subjects having a predisposition to
XX certain diseases associated with INR signaling. The present sequence is an
XX ISM protein described in the exemplification of the invention
XX
XX Sequence 352 AA;
XX
XX Query Match 15.9%; Score 373; DB 5; Length 352;
XX Best Local Similarity 29.8%; Pred. No. 1.1e-28;
XX Matches 104; Conservative 75; Mismatches 128; Indels 42; Gaps 13;
XX
XX 104 LRVSEAAIQISRCRV-SEYLNHLCYDLPDPSGKVIADINLPVKQSFHILHEQGIFVAP 162
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 29 LEFEDEAVESGCVYMRFMESHKCYDIVPTSSKLVVFDTLQVKKAFALVANGVRAAP 88
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 163 LWSDFRCQFVGLLSPDLDFILIRELETHGSLNTE-EQLETHITISAWKEAKROTGRNDSQ 221
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 89 LWESKKQSPVGMFLTITOFINILHRY--YKSPMVQIYELEEHKIEIETWRELYLQ-----ET 140
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 222 WRPQOHLVHATPYESLRDIAVKLLQNGISTVPVIYSSSDGSPFQLHLASLSGLKICIC 281
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 141 FKP---LVNTSPDASLDFAVYSLINKKHRLPVIDPISGNA-----LYLTHKILKFL- 191
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 282 RYFKNSTGNLPILNPQVCSIPLGSPVKPKIGDLSRPLAMLPNASSLSALMLVOAGVSS 341
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 192 QLEFSDMPKPAFMKQNLDELGIGTV-----HNIAFIHPTPIIKALNIFVERRISA 242
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 342 IPIVDDNDSLLDTSRSDITALAKQVYTHVRLDEMTHTQALQGDANTPFGEFNGQRC 401
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 243 LPVWDESGKVVDIYKEDVINLAAEKYNNL---DITVTQALQHRQ-----YFEG--V 291
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 402 QMCLRSPLKLMERLANPGVRRYFIV-EAGSKRVEGLISUSLIDIFKFL 449
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 292 VKCNKLEILTIIVIRAEVHRLVVVNEADS--IVGLISLSDILQALI 338
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX
XX RESULT 12
```

```
AA96793
ID AAY96793 standard; protein; 77 AA.
XX
XX AAY96793;
XX
XX 26-SEP-2000 (first entry)
XX
XX Wheat sucrose non-fermenting 4 protein from clone wreln.pk0143.62.
XX
XX Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
XX carbon catabolite repression; development; nitrogen partitioning.
XX
XX Triticum aestivum.
XX
XX WO200036116-A2.
XX
XX 22-JUN-2000.
XX
XX 15-DEC-1999; 99WO-US029825.
XX
XX 16-DEC-1998; 98US-0112564P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Heppard EP, Miao G, Weng Z;
XX
XX WPI; 2000-431594/37.
XX N-PSDB; AAA51357.
XX
XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
XX involved in carbon catabolite repression in plants and seeds, useful for
XX controlling carbon and nitrogen partitioning pathways during plant growth
XX and development.
XX
XX Claim 23; Page 48; 48pp; English.
XX
XX This is a partial wheat sucrose non-fermenting 4 (SNF4) protein which is
XX involved in carbon catabolite repression in plants and seeds. The cDNA
XX was isolated, based on similarity to SNF4 proteins from Arabidopsis
XX thaliana and Saccharomyces cerevisiae. The polynucleotides are used in
XX plants to control carbon and nitrogen partitioning pathways during plant
XX growth and development. The catabolite repression proteins would
XX facilitate studies for better understanding the mechanism of catabolite
XX repression in plants and could provide genetic tools to enhance or
XX otherwise alter the accumulation of carbohydrates, lipids and proteins
XX during plant growth and development
XX
XX Sequence 77 AA;
XX
XX Query Match 15.8%; Score 369; DB 3; Length 77;
XX Best Local Similarity 90.9%; Pred. No. 2.3e-29;
XX Matches 70; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
XX
XX 228 LVHATPYESLRDIAVKLLQNGISTVPVIYSSSDGSPFQLHLASLSGLKICICRYFKNS 287
XX 1 LVHATPYESLRGIAMKILETGISTVPIYSSSDGSPFQLHLASLSGLKICICRYFKNS 60
XX
XX 288 TGNLPILNPQVCSIPLG 304
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 61 TGSILNPQVCSIPLG 77
XX
XX RESULT 13
XX AAW29817
XX ID AAW29817 standard; protein; 331 AA.
XX
XX AAW29817;
XX
XX 16-MAR-1998 (first entry)
XX
XX Mammalian AMPK-gamma subunit protein.
XX
XX 5'-AMP activated protein kinase; AMPK; non-catalytic subunit;
```


KW protein phosphorylation; cholesterol; fatty acid; gamma subunit;
 KW hormone sensitive lipase; HSL; ss.
 XX Mammalia.

OS
 XX WO9725341-A1.
 PN
 XX 17-JUL-1997.

XX
 PF 07-JAN-1997; 97WO-US000270.
 XX
 PR 08-JAN-1996; 96AU-00007450.

XX (SVIN-) ST VINCENTS INST MEDICAL RES.
 XX (DART-) DARTMOUTH COLLEGE.
 XX

PI Kemp BE, Stapleton DI, Mitchelhill KI, Witters LA;
 XX WPI; 1997-372811/34.

XX New isolated 5'-AMP-activated protein kinase subunit(s) - used to develop
 PT products for treating e.g. hyper-cholesterolaemia, obesity, hypoxia,
 PT ischaemia, nutrition disorders or diabetes mellitus.

XX Claim 21; Page 54; 63pp; English.

XX This polynucleotide sequence encodes a mammalian 5'-AMP-activated protein
 CC kinase (AMPK) non-catalytic gamma subunit. AMPK polypeptides can be used
 CC to identify compounds which regulate the action of kinases. Such
 CC fragments can be used to reduce biosynthesis of cholesterol and fatty
 CC acids. They may also be used to inhibit the release of these molecules
 CC from intracellular stores by hormone sensitive lipase (HSL). They may
 CC also be used to reduce cellular malonyl CoA levels and promote the beta-
 CC oxidation of fatty acids by mitochondria. AMPK-alpha fragments could be
 CC used in the treatment of e.g. hypercholesterolaemia, hyperlipidaemia,
 CC obesity, clinical syndromes associated with hypoxia or ischaemia (e.g.
 CC myocardial infarction) disorders of nutrition and diabetes mellitus

XX Sequence 331 AA;

Query Match 15.7%; Score 367; DB 2; Length 331;

Best Local Similarity 29.1%; Pred. No. 4.1e-28;

Matches 111; Conservative 74; Mismatches 133; Indels 64; Gaps 14;

QY 69 LSPSTPGSRMNDVNFQRTVTLSDGTGTVSGTLRVSEAAIQISRCRVSEVNLHTCY 128
 DB 5 ISSDSPA-----VENEHPQETPE-SNNSV-----YTSPKSHRCY 39

QY 129 DLLPDSGKVALDINLPVQSFHILHEQGIPIVAPLWDSFGQVGLLSPDLFILIRELE 188
 DB 40 DLIPYSSKLWFDTSIQVKAFFALVTNGVRAAPLWDSKKQSFVGMLTITDFINILHRY - 98

QY 189 THGSNLTE-EQLETHTISAKKEAKQTNGENDSQWRPQCHLVHATPYESLRDIAVLQN 247
 DB 99 -YKSAUVQIYELEEHKTIETWREYVQ-----DSFKP---LVCISNAPLFDVAVSLRN 148

QY 248 GSTVPVYSSSDSGSPQLLHSLGILKICRYFQNSGTNLPIINQVCSIPFGSWV 307
 DB 149 KHLRPVIDPESGN-----TLYLTHKRIKFL-KLFTFFPKPFMSKSLLELQIGTYA 202

QY 308 PKIGDINSRPLAMLRPNASLSSALNMLVQGVSSIFVDNDLSLDTYSRSDITALAKDK 367
 DB 203 -----NIAMVRVTTTPVVALGIFVQHRVSGALPVDEKGRVDIYSKFDVINLAAEK 253

QY 368 VYTHVELDEWTHQALQODANTPPGFENGQRCQCLRSDDPLLKMYERLANPGRVFI 427
 DB 254 TNNL-----DVSVTKAQHRSN-----YFEG---VLKCYLHETLETIINRLVEAEVRLVW 302

QY 428 VEAGSKRVGGIISLSDFKFL 449
 DB 303 VDENDV-VKGIYSLSLDILQALV 323

RESULT 14

AAO18497
 ID AAO18497 standard; protein; 331 AA.

XX
 AC AAO18497;

XX
 DT 11-OCT-2002 (first entry)

XX Human insulin receptor signaling modifier SEQ ID NO: 16.

XX Human, insulin receptor signaling; insulin receptor signaling modifier;
 KW ISM; diabetes; metabolic syndrome; antidiabetic.

XX Homo sapiens.

XX WO200255664-A2.

XX 18-JUL-2002.

XX 11-JAN-2002; 2002WO-US001048.

XX 12-JAN-2001; 2001US-0261226P.

XX 12-JAN-2001; 2001US-0261303P.

XX 12-JAN-2001; 2001US-0261304P.

XX 12-JAN-2001; 2001US-0261335P.

XX 12-JAN-2001; 2001US-0261336P.

XX 12-JAN-2001; 2001US-0261361P.

XX 12-JAN-2001; 2001US-0261456P.

XX 12-JAN-2001; 2001US-0261457P.

XX 12-JAN-2001; 2001US-0261458P.

XX 12-JAN-2001; 2001US-0261459P.

XX 12-JAN-2001; 2001US-0261461P.

XX 12-JAN-2001; 2001US-0261518P.

XX 12-JAN-2001; 2001US-0261532P.

XX 12-JAN-2001; 2001US-0261589P.

XX 12-JAN-2001; 2001US-0261590P.

XX 12-JAN-2001; 2001US-0261694P.

XX 12-JAN-2001; 2001US-0261695P.

XX 12-JAN-2001; 2001US-0261697P.

XX (EXEL-) EXELIXIS INC.

XX Seidel-Dugan C, Ferguson KC, Kidd T;

XX WPI; 2002-599664/64.

XX N-PSDB; AAL48616.

XX Identifying an insulin receptor signaling modulator, useful as drug
 PT targets for treating diabetes or metabolic disorders, comprises
 PT contacting an assay system comprising insulin receptor signaling
 PT modifiers with a test agent.

XX Disclosure; Page 54-55; 232pp; English.

XX The present invention relates to a method of identifying a candidate
 CC insulin receptor (INR) signaling modulating agent, involving contacting
 CC an assay system comprising an insulin receptor signaling modifier (ISM)
 CC polypeptide or nucleic acid with a test agent, and detecting a test agent
 CC -biased activity of the assay system. The method is useful for
 CC identifying candidate INR signaling modulating agents. ISM genes may be
 CC used as drug targets for treatment of disorders related to INR signaling
 CC such as diabetes or metabolic syndrome. ISM nucleic acids and
 CC polypeptides are useful for identifying and testing agents that modulate
 CC ISM function and for other applications related to the involvement of ISM
 CC in INR signaling, and for identifying subjects having a predisposition to
 CC such diseases associated with INR signaling. The present sequence is an
 CC ISM protein described in the exemplification of the invention

XX Sequence 331 AA;

Query Match 15.7%; Score 367; DB 5; Length 331;

Best Local Similarity 29.1%; Pred. No. 4.1e-28;

Matches	111;	Conservative	74;	Mismatches	133;	Indels	64;	Gaps	14;
QY	69	LSSPSTPGSRMMNDVDNENFQRTVTLSDGTVSEGTLLFVSEAAIQISRCRVSEYLNLTCTY	128						
Db	5	ISSDSPA-----VENEHPQETPE-SNSNV-----YTSFMKSHRCY	39						
QY	129	DLSPDSGKVTALDINLPEVKOSFHLLHQQGTPVAPLWDSFRGQFVGLLSPDLDFILLIRELE	188						
Db	40	DLIPTSSKLVVFDTSLQVKKAFFALVTNGRAAPLWDSKQSFVGMLTITDIFNILHRY-	98						
QY	189	THGSNLTE-PQLTHHTISAMKEAKROTNGRNDSQWRPQQHLVHATPYESLRDIAVKLLON	247						
Db	99	-YXSALVQVIELEBHKLTETREWVLQ-----DSFKP---LVLCISPNASLFDVSSLRN	148						
QY	248	GISTVPVIYSSSDGSPQQLHLASLGLIKCIQRYFNKSTGNLPILNPQVCISPIPLGSWV	307						
Db	149	KIHLRPVIDESGN-----TLYILTHKRLKFL-KLFITFPKPEFNKSLSEELQIGTYA	202						
QY	308	PKTGDLSNRPLAMLRPNASLSALNMLVQAGVSSIPVDNDNSLLDTYRSRSDITALAKDK	367						
Db	203	-----NIAMVRITTPVYVALGIEFVQHRVSALPVVDEKGRVVDIYSKFDVINLAAEK	253						
QY	368	VYTHVRLDENTHQALQOLGQDANTPFGFNQRCQMCILRSDELKMKWERLANPGVRRVPI	427						
Db	254	TYNNL-----DVSVTVALQRRSH-----YFEG--VLKCYLHETLETIINRLVEAEVHRLVV	302						
QY	428	VEAGSKRVEGIIISLSDIFKFLL	449						
Db	303	VDENDV-VKGISVLSDLILQALV	323						
RESULT	15								
ID	ABU04258								
XX	ABU04258	standard; protein; 331 AA.							
AC	ABU04258;								
XX	XX								
DT	29-JAN-2003	(first entry)							
XX	XX								
DE	Human expressed protein tag (EPT) #924.								
XX	XX								
KW	Translational profiling; expressed protein tag; EPT; kinase; phosphatase;								
KW	protease; protease inhibitor; transporter; cytoskeletal protein;								
KW	receptor; transcription factor; cancer; MHC;								
KW	major histocompatibility complex; myeloma; colon cancer; gastric cancer;								
XX	adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.								
OS	Homo sapiens.								
XX	XX								
FN	WO200278524-A2.								
XX	XX								
PD	10-OCT-2002.								
XX	XX								
PF	28-MAR-2002; 2002WO-US009671.								
XX	XX								
PR	28-MAR-2001; 2001US-0279495P.								
XX	XX								
PR	21-MAY-2001; 2001US-0292544P.								
XX	XX								
PR	08-AUG-2001; 2001US-0310801P.								
XX	XX								
PR	01-OCT-2001; 2001US-0326370P.								
XX	XX								
PR	04-DEC-2001; 2001US-0336780P.								
XX	XX								
PR	20-FEB-2002; 2002US-0336985P.								
XX	XX								
PA	(ZYCO-) ZYCOS INC.								

Chicz RM, Tomlinson AJ, Urban RG;
WPI: 2003-040607/03.
New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

```

Example 2; SEQ ID NO 924; 134pp; English.

The invention describes a purified polypeptide, which comprises a
fragment of a kinase, phosphatase, protease, protease inhibitor,
transporter, cytoskeletal protein, receptor or transcription factor. The
polypeptide is useful as an immunogenic composition for eliciting in a
mammal an immunogenic response directed against any of the purified
polypeptide. The purified polypeptide, or the antibody that binds to this
polypeptide, is useful for treating cancer. The polypeptide is also
useful for identifying compounds that binds to a naturally processed
class I or class II MHC-binding polypeptide. The polypeptides and
polynucleotides are particularly useful for treating or preventing
myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
lymphoma or leukaemia. These are also useful for screening agents for
treating the above mentioned diseases. This sequence represents an
expressed protein tag (EPT) isolated from human tissue for translational
profiling. Note: This sequence does not appear in the printed
specification but was obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

XX      Sequence 331 AA;
SQ

Query Match          15.7%; Score 367; DB 6; Length 331;
Best Local Similarity 29.1%; Pred. No. 4.1e-28;
Matches 111; Conservative 74; Mismatches 133; Indels 64; Gaps 14;

QY      69  LSPSPTPGSRMNDVNEENFQRTVTLSDGTVSGTLRVSEAAIQISRCRVSEVLNLTQY 128
DB      5  ISSDSDSPA-----VENEHPQETPE--SNNSV-----VFSEMKSHRCY 39

QY      129  DLIPDSGKVIADINLVPQSFHILHEQGIPIVAPLMDSPRGQVGLSPDLFILRELE 188
DB      40  DLIPTSKLVVPTSLQVKKAFALVTNGVRAAPLMDSKQSFVGMLTITDFINILHRY- 98

QY      189  THGSNUTR-EOLETHYTISAWKEAKRQTNGRNDSCWRPQQHIVHATPYESLRDITAVKLLQN 247
DB      99  -YKSAALVQIVELPEHKIETWREVYLQ-----DSFKP---LVCISPNASLFDVSSLRN 148

QY      248  GISTVPVIYSSSDGSPFQLHLASLGLKCTCRVFNKSTGNLPIINQVPCSPILGWSV 307
DB      149  KHRLRPVDPESGN-----TLYITLTKRILKFL-KLFIETFPKPFMSKLEELQITYA 202

QY      308  PKIGDLNSRPLMLRPNASLSSALNMLQVAGVSSIPVDDNSDLDTPTYSRSDITALKDK 367
DB      203  -----NIAMVRTTPVVALGIFVQHRVSALPVDEKGRVVDIYKFDVINLAAEK 253

QY      368  VYTHVRLEDNTIHAQLQGDANTPPGFFNGQRCQMLRSDPLLKQWERLANPGVRVFI 427
DB      254  TYNKL---DVSVTKALQHRSH-----YFEG---VLKCYLHETLETTINRLVEAEVHRLDV 302

QY      428  VEAGSKRVGEHISLSDIFKPLL 449
DB      303  VDENDV-VKGVLSLSDILQALV 323

Search completed: July 7, 2004, 17:50:50
Job time : 61 secs

```

Search completed: July 7, 2004, 17:50:50
Job time : 61 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1030.5	44.0	391	2	B86222		hypothetical prote
2	355.5	15.2	330	2	T10759		AMP-activated prot
3	303.5	13.0	322	1	RGBYC3		regulatory protein
4	229.5	9.8	274	2	S67444		probable 5'-AMP-ac
5	228	9.7	478	2	T25899		hypothetical prote
6	154	6.6	460	2	T23248		hypothetical prote
7	123	5.3	270	2	T09514		5'-AMP-activated p
8	122.5	5.2	443	2	T25854		hypothetical prote
9	115	4.9	629	2	T18227		hypothetical prote
10	114	4.9	296	1	G64453		conserved hypothet
11	114	4.9	399	2	D96832		hypothetical prote
12	112.5	4.8	280	1	H64452		conserved hypothet
13	112.5	4.8	408	2	T43240		CBS-domain protein
14	108.5	4.6	168	1	A64478		hypothetical prote
15	108	4.6	1547	2	A13243		hypothetical prote
16	107.5	4.6	304	2	D71603		replication factor
17	107.5	4.6	4910	2	S64942		probable membrane
18	107	4.6	392	2	D71175		hypothetical prote
19	107	4.6	476	2	AB1294		glutamy1-tRNA (Gln)
20	106.5	4.5	567	2	D72359		conserved hypothet
21	106	4.5	476	2	AH1665		glutamy1-tRNA (Gln)
22	106	4.5	703	2	T43557		P-box/KD-repeat pr
23	106	4.5	842	2	C83177		probable phosphotr
24	106	4.5	1475	2	T33318		hypothetical prote
25	105	4.5	527	2	S64060		probable membrane
26	104.5	4.5	514	1	B69214		MJ0120 protein hom
27	104.5	4.5	781	2	B64135		ribonucleoside-dip
28	103.5	4.4	281	1	D69030		MJ1225-related pro
29	103.5	4.4	969	2	T02842		probable membrane

```
Qy 389 ANTFFGFFNGQRCQMLRSDSLKVMERLANP 421
: : : : : : : : : : : : : : : : : :
Db 354 ASPPYGIFNGQRCMLRSDSLKVMERLANP 386

RESULT 2
T10759
AMP-activated protein kinase (EC 2.7.1.-) gamma chain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T10759
R;Woods, A.; Cheung P, C.F.; Smith, F.C.; Davison, M.D.; Scott, J.; Beri, R.K.; Carling,
J. Biol. Chem. 271, 15282-15290, 1996
A;Title: Characterization of AMP-activated protein kinase beta and gamma subunits: Assen
A;Reference number: Z56738; MUID:96215327; PMID:8626596
A;Accession: T10759
A;Status: translated from GE/EMBL/DBD
A;Molecule type: mRNA
A;Residues: 1-330 <WOC>
A;Cross-references: EMBL:X95578; NID:gl185270; PIDN:CAA64831.1; PID:gl185271
A;Experimental source: strain Wistar
C;Complex: heterotrimer; alpha, beta and gamma chains
C;Function:
A;Description: is responsible for the regulation of fatty acid synthesis by phosphorylat
C;Superfamily: CAT3 protein
C;Keywords: fatty acid biosynthesis; phosphotransferase

Query Match 15.2%; Score 355.5; DB 2; Length 330;
Best Local Similarity 30.3%; Pred. No. 6.5e-20;
Matches 100; Conservative 67; Mismatches 124; Indels 33; Gaps 11;

Qy 121 YLNHTCYDLPDSGKVIADINLPVKQSFHLHQGIPVAPLWDSRFGQVGLLSP 180
: : : : : : : : : : : : : : : : : :
Db 31 FMSKSHRCYDLIPTSSKLVFDTLSQVKKAFALVTNGVRAAPLWDSKQSFVGMLTITDF 90
: : : : : : : : : : : : : : : : : :

Qy 181 ILILRELETHGCSNLTE-EQLETHTISAWKEAKQTNGRNDQWRPQOHLVHATPYESLRD 239
: : : : : : : : : : : : : : : : : :
Db 91 INILHRY--YKSAUVQYVEEHEKLETWREVYLO-----DSFKP--LVCSINASLFD 139
: : : : : : : : : : : : : : : : : :

Qy 240 IAVKLQNGISTVPVIYSSSDSGSFPPQLHLASLSGLKICRYFKNSGTNLPILNOPVC 299
: : : : : : : : : : : : : : : : : :
Db 140 AVSSLIRKTHRLPVIDPESGN-----TLYLTHTKRLKFL-KLFTIEFPKPEWKSLE 193
: : : : : : : : : : : : : : : : : :

Qy 300 SIPLGSWVPKIGDLSRPLAMLRPNASLSSALNMLVQAGVSSIPVDNDLSLDTYGRSD 359
: : : : : : : : : : : : : : : : : :
Db 194 ELQIGTYA-----NIAMVRTTTPYVALGIFVQHRVSALEPVDVEKGRVVDIYKFD 244
: : : : : : : : : : : : : : : : : :

Qy 360 ITALAKDKVYTHVELDEMTIHOALQLGQDANTPPGFNGQRCQMLRSDPILKVMERLAN 419
: : : : : : : : : : : : : : : : : :
Db 245 VINLAEEKTYNNL---DVSVTKALQHRSH-----YFEG---VLKCYLHETLEATINLVE 293
: : : : : : : : : : : : : : : : : :

Qy 420 PGRVRVFIVEAGSKRVEGIISLSDFKFL 449
: : : : : : : : : : : : : : : : : :
Db 294 AEVRLVVDDEHDV-VKGIIVSLDILQALV 322
: : : : : : : : : : : : : : : : : :

RESULT 3
RBYCY3
regulatory protein SNF4 - yeast (Saccharomyces cerevisiae)
N;Alternate names: CAT3 protein; protein G2945; protein YGL115W
C;Species: Saccharomyces cerevisiae
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jul-2000
C;Accession: A38906; J03016; S48508; S64125
R;Celienza, J.L.; Eng, F.J.; Carlson, M.
Mol. Cell. Biol. 9, 5045-5054, 1989
A;Title: Molecular analysis of the SNF4 gene of Saccharomyces cerevisiae: evidence for p
A;Reference number: A33480; MUID:90097921; PMID:2481228
A;Accession: A38906
A;Molecule type: DNA
A;Residues: 1-322 <CEL>
A;Cross-references: GB:M30470; NID:gl172635; PIDN:AAA35061.1; PID:gl172636
R;Schueller, H.J.; Entian, K.D.
Probable 5'-AMP-activated, gamma subunit family - fission yeast (Schizosaccharomyces p
C;Species: Schizosaccharomyces pombe
C;Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: T38059; S67444
R;McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z21766
A;Accession: T38059
A;Status: preliminary
A;Molecule type: DNA

Gene 67, 247-257, 1988
A;Title: Molecular characterization of yeast regulatory gene CAT3 necessary for glucos
A;Reference number: J03016; MUID:89006284; PMID:3049255
A;Accession: J03016
A;Molecule type: DNA
A;Residues: 1-322 <SCH>
A;Cross-references: GB:M21760; NID:gl171164; PIDN:AAA34472.1; PID:gl171165
R;Doi, A.; Doi, K.
submitted to the EMBL Data Library, June 1993
A;Description: Correct end of the ORF for the CDC20 gene of Saccharomyces cerevisiae.
A;Accession: S48507
A;Reference number: S48508
A;Molecule type: DNA
A;Residues: 1-21 <DOI>
A;Cross-references: EMBL:D16506; NID:G391938; PIDN:BA003958.1; PID:g2160324
R;Lauquin, G.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64122
A;Accession: S64125
A;Molecule type: DNA
A;Residues: 1-322 <LAU>
A;Cross-references: EMBL:Z72637; NID:gl322666; PIDN:CAA96823.1; PID:gl322667; GSPDB:GN
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:SNF4; CAT34; MIPS:YGL115W
A;Cross-references: SGD:S0003083; MIPS:YGL115W
A;Map position: 7L
C;Function:
A;Description: involved in derepression of glucose-repressed genes
C;Superfamily: CAT3 protein
C;Keywords: nucleus; transcription regulation

Query Match 13.0%; Score 303.5; DB 1; Length 322;
Best Local Similarity 27.9%; Pred. No. 6.8e-16;
Matches 95; Conservative 64; Mismatches 123; Indels 53; Gaps 12;

Qy 118 VSEYLNHTCYDLPDSGKVIADINLPVKQSFHLHQGIPVAPLWDSRFGQVGLLSP 177
: : : : : : : : : : : : : : : : : :
Db 22 IRKFLNKTSDYVLPVSVYRLVDTLSLVKSLNVLQNSIVSAPLWDSKTSRFGALLTT 81
: : : : : : : : : : : : : : : : : :

Qy 178 LDFILILRELETHGCSNLTE-EQLETHTISAWKEAKQTNGRNDQWRPQOHLVHATPY 234
: : : : : : : : : : : : : : : : : :
Db 82 TDFINV---IQYFSPNDKPELVD-----KLQDGLKDIERALGVQDQDTSIHPS 129
: : : : : : : : : : : : : : : : : :

Qy 235 ESLRDIAVLQNGISTVPVIYSSSDSGSFPPQLHLASLSGLKICI---CR---YFKNST 288
: : : : : : : : : : : : : : : : : :
Db 130 RPLFEACLKMLESRSGRIPLI-DQDEETHREIVVSVLTQYRILKFPVALNCRETHFLKIP 188
: : : : : : : : : : : : : : : : : :

Qy 289 GNLPILNQPVCSIFLGSWVPKIGDLSRPLAMLRPNASLSSALNMLVQAGVSSIPVDN 348
: : : : : : : : : : : : : : : : : :
Db 189 GDLNITQD-----NMKSCQMTTP---VIDVIOMLTQGRVSSVPIIDEN 229
: : : : : : : : : : : : : : : : : :

Qy 349 DSLDTYSRSDITALAKDKVYTHVELDEMTIHOALQLGQDANTPPGFNGQRCQMLRSD 408
: : : : : : : : : : : : : : : : : :
Db 230 GYLINVEAYDVLGKGIYNDLSL---SVGEALMRSD-----FEG---VYTCTKND 278
: : : : : : : : : : : : : : : : : :

Qy 409 PLLKVMERLANPGVRVFIVEAGSKRVEGIISLSDFKFL 449
: : : : : : : : : : : : : : : : : :
Db 279 KLSTIMDNIRKARVHRFFVVD-DVGRLVGVLTLSDLKYL 318
: : : : : : : : : : : : : : : : : :

RESULT 4
S67444
Probable 5'-AMP-activated, gamma subunit family - fission yeast (Schizosaccharomyces p
C;Species: Schizosaccharomyces pombe
C;Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: T38059; S67444
R;McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z21766
A;Accession: T38059
A;Status: preliminary
A;Molecule type: DNA
```

RESULT 6
T24248
hypothetical protein R53.7 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24248
R:Wilkinson, J
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19863
A:Accession: T24248
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-460 <WIL>
A:Cross-references: EMBL:Z66515; PFTN:CAA91351.1; GSPDB:GNO0020; CESP:R53.7
A:Experimental source: clone R53
C:Genetics:
A:Gene: CESP:R53.7
A:Map position: 2
A:Introns: 30/1; 133/1; 407/3

[illegible]

```

358 SDI-TALAKDKVYTHVRLDEMTIHQALQLQGDANTPFQFNGQRCQMCILRS-DPLLKYME 415
      :||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
344 TDICKALPRNFIEPKWLOETKVSUILHI-----CKSQILISSADSVGGVLD 390
      :||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      :||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
416 BLANPGVRRVPIVEAGSKRVEGIISLSDIFKLL 449
      :||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
391 TLLAGDTQSAPATHNG--KAIGVISLDTFLSHIL 422
      :||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      :||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 7
T09514
5'-AMP-activated protein kinase (EC 2.7.1.-) beta-1 chain - human
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T09514
R:Stapleton, D.; Woollatt, E.; Mitchelhill, K.; Nisholl, J.K.; Fernandez, C.S. et al.
FEBS Lett. 409, 452-456, 1997
A:Title: AMP-activated protein kinase isoenzyme family: subunit structure and
A:Reference number: Z16707; MUID:97367941; PMID:9224708
A:Accession: T09514

```

```

A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-270 <STA>
A:Cross-references: EMBL:Y12556; NID:g2230862; PIDN:CAA73146.1; PID:g2230863
C:Genetics:
A:Map position: 12
C:Complex: heterotrimer; alpha, beta and gamma chains
C:Function:
A>Description: is responsible for acting as a metabolic sensor for AMP levels
C:Keywords: fatty acid biosynthesis; heterotrimer; phosphotransferase

Query Match      5.3%; Score 123; DB 2; Length 270;
Best Local Similarity 34.7%; Pred.No. 0.053;
Matches 26; Conservative 13; Mismatches 20; Indels 16; Gaps 2;

QY   17 FQAICSLSPGIHYKFFVDGEWRHDERQPTISGEFGIVNTLYL-----TR 61
Db    |||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
112 FVAILDLPEGHQHKFVDCQWTHDSEPIVTSQLGTANNIQCWKXTDFEFDALMVDSQ 171
QY   62 EYNQINTLSSPSTG 76
Db    :::||:|||||:::|||||:::|||||:::|||||:::|||||:::
172 KCDVSSELSS-SPPG 185

RESULT 8
T25854
hypothetical protein T01B6.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T25854
R:Johnson, D.
submitted to the EMBL Data Library, August 1996
A>Description: The sequence of C. elegans cosmid T01B6.
A:Reference number: Z20:00
A:Accession: T25854
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-443 <JOH>
A:Cross-references: EMBL:U67950; PIDN:AAH07568.1; GSPDB:GN00028; CBSP:T01B6.3
A:Experimental source: strain Bristol N2; clone T01B6
C:Genetics:
A:Gene: CBSP-T01B6.3
A:Map position: X
A:A:introns: 5/3; 45/2; 72/1; 105/3; 153/3; 179/2; 200/3; 225/3; 255/3; 349/3; 397/2

Query Match      5.2%; Score 122.5; DB 2; Length 443;
Best Local Similarity 19.9%; Pred.No. 0.12;
Matches 77; Conservative 71; Mismatches 115; Indels 123; Gaps 21;

QY   112 QTSRCRVS-----EYNILHTCYDLLDPDSKVIALDNLPVKOSFHILHQG--- 157
Db    |||:||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6C QRARCTAAKIQRHADPPYHTPMKSITCYDLOPTHSSLVFDGKTAKVAHVALSQHGI 119
QY   158 IPVAPLWDSPRGQGV-GLLSPDLFILLR--ELETHGSNLTE--SOLETHITSAMKEAKR 212
Db    |||:||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
120 AAVVTMDTKQAECVFNMGCHLTAILLVAGNRREVASKTLVEFLKEIGSGNIIC---- 173
QY   213 QTNGRNDSSRRPOQHVLVHATPYESLRDIAYKLQNGISTVPV-----IYGSSSD 261
Db    |||:||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
174 ---SGVNQSVWEANAIIISH-----NKISFPPIFTTIIPKPGTPLYELT-- 213
QY   262 GSPPQLL---HLASLSGLKCICTRYFKNSKGNIPLINQPVCSTPLGSM---VPKIGDLNS 315
Db    |||:||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
214 ----PRMILQETV:KLSDFGDAILLVRQAT-----LDQX-----KIGTWDDVLKIG---- 257
QY   316 RELAMLRPNASLSSALNMVLQVGVSSIPVDDNDSLDTYSRSDFITALAKDKVYTH---- 371
Db    |||:||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
258 -----LNTTIEEAIKLASERKWSIPVNDPFQVINMLARKDIIIL----EIMSHQGNG 306
QY   372 ----VRIDEMTIHQALO----LGQDANTPFGFFNQGRQCMLSRDPLIKVM--ERLANPGV 422
Db    |||:||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
307 FHDMLEKPVKILQSLOSRLVGRSSYYTF-----ETVAQMNTSDKSLP--- 350

```


Qy 309 KIGDILNSRPLAMLRPNASLSSALNMLVQAGVSSIPVDNDLSLDITYSRSDITALAKS-- 366
|::: : : : |::: : : |::: : : |::: : : |::: : :
Db 171 KYGDVGKIVKEWTTNPNCITRETAKLFPAKYISGAPFV-DNDKLVGVISLHDAEINDID 229
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Qy 367 -KVYTHVRLEDMITHQALQLQGQANTFFGFENGQRCCMCLESDPLLKVMERLANFGVRV 425
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 230 KKVKVEMRRRDVITHKDEKI-YDA-----LTKMNMK---NRVGL 264
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Qy 426 FIVEAGSKRVEGHISLSIDIEKFL 448
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 265 VIVDDNNK-IVGIITRTDIKKII 286
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 11

D96832
hypothetical protein F18B13_17 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96832
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizlar, L.
Nature 433, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Iruos, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MOID:21016719; PMID:11130712
A:Accession: D96832
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <STO>
A:Cross-references: GB:AE005173; NID:g5902384; PIDN:AAD55486.1; GSPDB:GM00141
C:Genetics:
A:Gene: F18B13_17
A:Map position: 1

```

Query Match      4.9%; Score 114; DB 2; Length 399;
Best Local Similarity 19.3%; Pred. No. 3.47;
Matches 87; Conservative 86; Mismatches 151; Indels 126; Gaps 21;

QY      64 NCINTLSPSPGSRQMDVDNENFQRTVTITDGTWSEGLTVSEAAIQI---SRCSVSE 120
      3 NEVTTKASK---QRELVHKRPPIFFETLTIQ-----IQIFMQSYSLMQ 44
QY      121 YLNLHTCYDLLPDSGKVIALDINLPVQSFHLHQEQ--PVA-----PL---W--- 164
      45 F---KVKDLMDIKRRRLVEVPDNNATLGDALNTMTILGKPVANVRVRAVPAAKPGQWLGA 100
QY      165 -----DSFRGQFVGLLSPFLILLIRELTHGSMNLTQEQLETHTISAWKEAKR 212
      101 GGSMIVELDKSGSRARQYIGWMTLDVV-----AHAGDGE----- 138
QY      213 QTNGRNDQWRPOOHLV-----HATPYESLRDIAVKLLQMGISTVPEVIYSSSDG 262
      139 --SGLDKKMAAPVSIIGHCPEGISLWSLNNTSIMG--CMEMLSKGTHRVLPVLDNSTN 195
QY      263 -SEPOLLAHLASLGILX---CICRYEKNSTGNLPIINQPVCSIFLGSWVPKIGDLSRPL 318
      196 ITGHELVEASAYANLSQMDLISFFFPQSQOHLGILSHVTVD-----LSAIHNVIL 246
QY      319 AMLRPNASLGSALNMLVQAGVSSPIV-----DENDSLLD-----TVSRSDITALA 364
      247 A-LTSQARXKDAIQCMGIAMLNVAPIVEASGEGEDHQLVDGKNRRVVVGTFASDLKGCH 305
QY      365 KDKVYTHRWLDDEM---TIHQALQLQGDANTPPGFQGRQCOMCLRSDPLLKVMERLANP 420
      306 LATURSLPLNALSEVKEIKPRTLFTTAATSTP-----GRELVTHVTTSLAQVHHMVTTK 360
QY      421 GVRRVFTVEAGSKRVEGIIISLSDIFKELLS 450

```

Db 361 RVHRVWVWDQNG-GLQGLVSLTDIIAVVRS 389

RESULT 12

H64452

conserved hypothetical protein MJ1225 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 10-Sep-1999 #sequence_revision 13-Sep-1999 #text_change 21-Jul-2000

C:Accession: H64452

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.M.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Frazer, C.M.; Smith, H.O.; Woese, A.; Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: H64452

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-280 <BUL>

A:Cross-references: GB:U67563; GB:L77117; NID:g2826379; PIDN:AAB99228.1; PID:gl591856;

C:Genetics:

A:Map position: FOR1167099-1167941

C:Superfamily: conserved hypothetical protein MJ1225; CBS homology F:157-204/Domain: CBS homology <CBS>

Query Match 4.8%; Score 112.5; DB 1; Length 280;

Best Local Similarity 20.6%; Pred. No. 0.37;

Matches 46; Conservative 41; Mismatches 101; Indels 35; Gaps 6;

Qy 228 LVHATPPYBSLRDIAVKLLQNGISTVPVIYSSSD--GSPFQLLELASGLKLCICRYPK 285

Db 14 IVTVPTTIRKALMTNENKVRLLPVNAGNKKVVGIIITSNDIVDFMGSGSKYNLIREK 73

Qy 286 NSTGNLPILNPWCSPIGSWVPKTDGINSRPLAMLRPNASLSSALNMLVQAGVSSIPIV 345

Db 74 HERNFLAAINPV-----REIMEENVITIKENADIDEATETFLTNNVGAPIV 121

Qy 346 DDNSLDITYSRSDITALKDKVYTHVSLDEMTIHQALQLGQDANTPFGEFGRCQMCL 405

Db 122 NDNQLISLITERVIRALLDKIDENEVIDDYITRDVI-----VAIP-----GER----- 166

Qy 406 RSDPLLKVMERLANPGVRRVFIVEAGSKRVGGIISLSDIFKFL 448

Db 167 ---LKDVARTVWRNGFRRLPVWSEG--RLVGIIITSTDFIKLL 203

RESULT 13

T43240

CBS-domain protein Sds23p - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C:Accession: T43240; T40589; T43242

R:Ishii, K.; Kumada, K.; Toda, T.; Yanagida, M. EMBO J. 15, 6629-6640, 1996

A:Title: Requirement of Ppl phosphatase and 20S cyclosome/APC for the onset of anaphase

A:Reference number: Z22359; MUID:97133292; PMID:8978689

A:Accession: T43240

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-408 <ISH>

A:Cross-references: EMBL:D86840; NID:gl486256; PIDN:BAAL13172.1; PID:gl486257

R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, January 1999

A:Reference number: Z21938

A:Accession: T40589

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-408 <SSE>

A:Cross-references: EMBL:AL035216; PIDN:CAA22817.1; GSPDB:GN00067; SPDB:SPBC646.13

A:Experimental source: strain 972h-; cosmid c646

R:Kawamukai, M.

Biochim. Biophys. Acta 1446, 93-101, 1999
 A>Title: Isolation of a novel gene, moc2, encoding a putative RNA helicase as a suppressor of the growth defect of a yeast mutant
 A:Reference number: Z22361; MUID:99326140; PMID:10395922
 A:Accession: T43242
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-408 <RAW>
 A:Cross-references: EMBL:D87870; NID:g1565198; PIDN:BAAL3486.1; PID:g1565199
 C:Genetics: <SEE1>
 A:Gene: SPBC646.13
 A:Map position: 2
 A>Note: sds23

Query Match 4.8%; Score 112.5; DB 2; Length 408;
 Best Local Similarity 20.6%; Pred. No. 0.64;
 Matches 75; Conservative 57; Mismatches 121; Indels 111; Gaps 16;
 QY 156 QGTPVAPLWDSFRGQFVGLLSP-----LDFILLIRELE-----THGSN----- 193
 DB 43 QDIPVTFLDN-----ETALIDPETSMEERASSILLDRLSALPIVAAGSNEIATTFDYAD 98
 QY 194 -----LTEEQLETHITISAWKEAKRQTRGRNDSSQWRPQOHLVHAT 232
 DB 99 LNSFLMWVGFDDFNDGRFKKVAEDIRAGKVITAVEAKL--GKNKDDFIT--IPHTT 152
 QY 233 PYESLSDIAVLLQNGISTVPVIYSSSDGSPFOLLHLASLSGLIKICICRYFNKSTGNLP 292
 DB 153 SLGLRLAEI-----LSSGIRRVAV-----TNEQG-----ELSFMSQRSII-----RFLNNIRAFP 198
 QY 293 ILNQPVCSIFLGSWVPKIGDLSRPLAMRPNASLSLALNMLVQAGVSSIPVINDNSLL 352
 DB 199 DL-EPLMSRIHSL-----DIGSTDITCISGDQKVAALQRMQNTGIGSLAVVDAQFRL 252
 QY 353 DTYSRSDITALAK-----DKVYTHVRLDEMTIHQALQIQDQANTPFPGFNGQRCOMCL 405
 DB 253 GNISLVNDVKVYTRSSSVYLLNKSCHAF-LSVIKSEQIGIRAGKLSAPAFNIY----- 302
 QY 406 RSDPLKLMERLANPGVRRVFIIVEAGS-----KRVGIIISLSDI 444
 DB 303 ESSTFAFTLAKVATQCHRLWLQVSPSCPPSPKNAHLSPSGMGVGVKNOLLGVVSLTDI 362
 QY 445 FKFL 448
 DB 363 ISVL 366

RESULT 14
 A64478
 hypothetical protein homolog MJ1426 - Methanococcus jannaschii
 C:Species: Methanococcus jannaschii
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: A64478
 R:Bult, C.J.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Bult, C.J.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Rison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1258-1073, 1996
 A:Authors: Kaire, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: A64478
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-168 <BL>
 A:Cross-references: GB:U67583; GB:L77117; NID:g2826409; PIDN:AAB99437.1; PID:g1592076; T43242
 C:Genetics: FOR1396048-1396554
 A:Map position: TTG
 A:Start codon: TTG
 C:Superfamily: conserved hypothetical protein yhcV; CBS homology
 F:120-168/Domain: CBS homology <CBS>

Query Match 4.6%; Score 108.5; DB 1; Length 168;
 Best Local Similarity 26.2%; Pred. No. 0.35;
 Matches 38; Conservative 32; Mismatches 48; Indels 27; Gaps 7;

QY 234 YE--SLRDIATVLLQNGISTVPVIYSSSDGSPFOLLHLASLSGLIKICICRYFNKSTGNL 291
 DB 28 YEDNDLIDVIRLFRKNKISGAPVL--NKDG--KLWGIISSEDIIVTIIVTH--NEDLNL 79
 QY 292 PILNQP--VCSIPL-----GSWVPKIGDLSRPLAMRPNASLSLALNMLV 335
 DB 80 -ILPSPDLIELPKTALKIEEFMEDLKNALKTKVRDVMTRKVIVAKPDMTINDAAKLWV 138
 QY 336 QAGVSSIPVDDNDNSLLDTYRSDI 360
 DB 139 KNNIKRLPVVDDEGLIGIVTRGDL 163
 RESULT 15
 A12843
 hypothetical protein alr1903 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change C9-Dec-2002
 C:Accession: A12843
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguc, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 205-213, 2001
 A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: A12843
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1547 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAH73602.1; PID:g17130993; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr1903

Query Match 4.6%; Score 108; DB 2; Length 1547;
 Best Local Similarity 21.0%; Pred. No. 10;
 Matches 81; Conservative 62; Mismatches 173; Indels 70; Gaps 15;
 QY 98 TVSEGTURVSEAAIOISRCV-----SEYLNLTCTYDILLPSGKVIALDINKLPVQSPHIL 153
 DB 48 TVQSAKEERSASQOVERLDVLAGLRKYAOEHVLLVGRPGSGKSTAL--LRLLEAEERPL 106
 QY 154 -----HEQIPVAPLWDSFRGQFVGLLSPDLDFILIRELTHGNSLITEQLETHITISAW 207
 DB 107 POPLDAERGVLAPP--SLAGKGVGGIPIVLVELRYTQTSILDLIRDFLKHCLLD 163
 QY 208 KEAKRQTRNGRNDSSQWRPQOHLVHATPYESLRFDIKVLQNGISTVPVIYSS-----SSDG 262
 DB 164 TATIEQL--LFDGQFLLVDGINELPSEKARQDLKFRQDNQKTPMIFTTRDLGVGGDL 221
 QY 263 SFPOLLHLASLSG--ILKICRYEK-----NSTGNLPILNQPVCS--SIFLG 304
 DB 222 GITKLEMQPLTAEQMQQFVRAVLPAGQEQMLQNLGDRLREFGETPILLFMLCLLFFVETG 281
 QY 305 SWVPKIG--DLNSRPLAMRPNASLS-----ALNMLVQAGVSSIPVDD 347
 DB 282 DIPNLGLLFRQLTQRYKNELKPNVPKGVSRDHSWMLQLAFNM-----TKGDKLTEL 336
 QY 348 NDSLDTYSSRDTALAKDKVYTHV-----LDPMTHQALQIQDQANTPFPGFNGQRC 401
 DB 337 NVAILKSKAEVITQLLNKDEGNFPRDQASELSKDLNHLIQLA--ANDQIE-FRHQLI 393
 QY 402 QMCLRSDDLKVMERLANPGVRRVFI 427
 DB 394 QEYTTAECLLKLPSLSDSLQREYL 419

Search completed: July 7, 2004, 17:52:51
 Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:47:08 ; Search time 17 Seconds
(without alignments)

1381.391 Million cell updates/sec

Title: US-09-857-525C-2

Perfect score: 2342

Sequence: 1 TREHLPMSPIEGPTVFQAI.....SKRVEGISLSDFKLLSL 451

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	373	15.9	569	1 AAKH HUMAN	Q9ugj0 homo sapien
2	367	15.7	331	1 AAKG HUMAN	P54619 homo sapien
3	365	15.6	330	1 AAKG BOVIN	P58108 bos taurus
4	355.5	15.2	330	1 AAKG RAT	P80385 rattus norv
5	350	14.9	330	1 AAKG MOUSE	O54950 mus musculus
6	340	14.5	464	1 AAKI HUMAN	Q9ugi9 homo sapien
7	334.5	14.3	514	1 AAKI PIG	Q9myr4 sus scrofa
8	303.5	13.0	322	1 SNF4 YEAST	P12904 saccharomyc
9	289	12.3	334	1 YL28 SCHPO	Q10343 schizosacch
10	261.5	11.2	328	1 SNF4 KLULA	Q9p869 kluyveromyc
11	124	5.3	269	1 AAKB MOUSE	Q9r078 mus musculus
12	124	5.3	269	1 AAKB RAT	P80386 rattus norv
13	123	5.3	269	1 AAKB HUMAN	Q9y478 homo sapien
14	122.5	5.2	271	1 AAKB RAT	Q9qz44 rattus norv
15	121	5.2	272	1 AAKC HUMAN	Q43741 homo sapien
16	116	5.0	122	1 AAKB PIG	P80387 sus scrofa
17	114	4.9	296	1 YC32 METJA	O58629 methanococc
18	112.5	4.8	280	1 YC25 METJA	O58622 methanococc
19	112.5	4.8	408	1 SD23 SCHPO	O09826 schizosacch
20	108.5	4.6	168	1 YE26 METJA	Q58821 methanococc
21	107.5	4.6	4913	1 MDN1 YEAST	Q12019 saccharomyc
22	107	4.6	476	1 GATE LISMO	Q8y6d3 listeria mo
23	106.5	4.5	548	1 PPAC THEMA	Q9wz56 thermotoga
24	106	4.5	476	1 GPAT LISIN	Q92a04 listeria in
25	106	4.5	703	1 POP2 SCHPO	O14170 schizosacch
26	105	4.5	527	1 YGFC YEAST	P43172 saccharomyc
27	104.5	4.5	756	1 RIRI HAEIN	P43754 haemophilus
28	101.5	4.3	302	1 YR33 THEPE	P15889 thermofilum
29	101	4.3	509	1 Y100 METJA	Q57564 methanococc
30	100.5	4.3	958	1 MSH4 MOUSE	O99mt2 mus musculus
31	99.5	4.2	1024	1 SRE1 RAT	P56720 rattus norv
32	98.5	4.2	3343	1 YOG7 CABEL	P34616 caenorhabdi
33	97.5	4.2	238	1 YD43 HAEIN	P71379 haemophilus

RESULT 1

ID	AAKH HUMAN	STANDARD;	PRT;	569 AA.
AC	Q9UGJ0: Q9NUZ9; Q9UDN8; Q9ULX8;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	5'-AMP-activated protein kinase, gamma-2 subunit (AMPK gamma-2 chain)			
DE	(AMPK gamma2) (H91620p).			
GN	PRKAG2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RX	SEQUENCE FROM N.A. (ISOFORM A).			
RA	MEDLINE=20164049; PubMed=10698692;			
RA	Cheung P.C.F., Salt I.P., Davies S.P., Hardie J.G., Carling D.;			
RT	*Characterization of AMP-activated protein kinase gamma-subunit			
RT	isoforms and their role in AMP binding.";			
RL	Biochem. J. 346:659-669 (2000).			
RN	[2]			
RX	SEQUENCE FROM N.A. (ISOFORM B).			
RA	MEDLINE=20564210; PubMed=1112354;			
RA	Lang T., Yu L., Qiang T., Jiang J., Chen Z., Xin Y., Liu G., Zhao S.;			
RT	*Molecular cloning, genomic organization, and mapping of PRKAG2, a			
RT	heart abundant gamma-2 subunit of 5'-AMP-activated protein kinase, to			
RL	human chromosome 7q36.";			
RL	Genomics 70:258-263 (2000).			
RN	[3]			
RX	SEQUENCE FROM N.A. (ISOFORM B).			
RC	TISSUE=Placenta;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Orsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,			
RA	Matsumura H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,			
RA	Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;			
RT	*NEDO human cDNA sequencing project.";			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBSJ databases.			
RN	[4]			
RX	SEQUENCE FROM N.A. (ISOFORM B).			
RC	TISSUE=Liver;			
RA	MEDLINE=22398257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullay S.J.,			
RA	Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			

34	97.5	4.2	701	1 YL66 YEAST	Q06149 saccharomyc
35	97.5	4.2	2052	1 FYV1 MOUSE	Q9z1t6 mus musculus
36	97	4.1	196	1 Y525 METKA	P50100 methanopyru
37	97	4.1	909	1 NUOG SHEON	Q8ei34 shewanella
38	96	4.1	551	1 IL2B HUMAN	P14784 homo sapien
39	96	4.1	2037	1 FASI CANAL	P34731 c fatty aci
40	95.5	4.1	656	1 MUTL LACIA	Q9cd11 lactococcus
41	95.5	4.1	977	1 YAG9 SCHPO	Q09872 schizosacch
42	95	4.1	690	1 NCPR PHAHO	P37116 phaseolus a
43	94	4.0	465	1 YGTA YEAST	P53101 saccharomyc
44	94	4.0	527	1 YB64 YEAST	P38314 saccharomyc
45	94	4.0	951	1 LGR4 RAT	Q9z2h4 rattus norv

ALIGNMENTS


```
Db 246 LEFEDEAVESGVYRMFRSHKCYDIVPTSSKLVFDTTLQVKKAFVALVANGVRAAP 305
Qy 163 LWDSPRGQVGLSPDLFILRELETHGSLNLE-BOLETHHTISAMKEAKRQINGRNDQ 221
Db 306 LWSKKQSFVGMUTITDFINLHRY--YKSPWQIYELEHETWRELYLQ-----ET 357
Qy 222 WRPQOHLVHATPYESURDIAVKLQNGISTVPVIYSSSDGSPQLHLASLGLKLCIC 281
Db 358 KXP---LVNISPDASLFDVAVSLIKKHRLPVIDPISGNA-----LYILTHKRLKFL- 408
Qy 282 RFPKSTGNLPIINQVPCSPILGSPWPKIGDLNRLPLMRPNASLSSALNMLVQGVSS 341
Db 409 QLFMSDMPKPAFMKQNLDELIGTY-----HNIAFTHTPTPIKALNIFVRRISA 459
Qy 342 IPTVDNDSLLTYSRSDITALAKQVYTHVRLEDEMTHQALQGLQGDAN--PFGFNGQRC 401
Db 460 LPVVDSESGKVVDIYSPFEDVINLAETYNL-----DITVTQALQHSQ-----YFEG--V 508
Qy 402 QMCIRSDPILKVMERLANPQVRRFVIV-EAGSKRVEGIISLSDFIKFL 449
Db 509 VXCNKLEILTIIVDRIVRAEVRHLVVVNEADS--IVGIISLSLQALQALI 555

RESULT 2
AAKG HUMAN
ID AAKG HUMAN STANDARD; PRT; 331 AA.
AC P54619;
DT 01-OCT-1996 (Rel. 34, Created)
DZ 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE (AMPKg).
GN PRKAG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Petal liver;
RX MEDLINE=96224074; PubMed=8621499;
RA Gao G., Fernandez C.S., Stapleton D., Auster A.S., Widmer J.,
RA Dyck J.R.B., Kemp B.E., Witters L.A.;
RT "Non-catalytic beta- and gamma-subunit isoforms of the 5'-AMP-
RT activated protein kinase."
RL J. Biol. Chem. 271:8675-8681(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares K.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.C., Hulyk S.W.,
RA Villalón D.K., Murty D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Scherch A., Schein J.E., Jones S.D.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
```

```
CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U42412; AAC50495.1; -.
DR EMBL; BC000358; AAH00358.1; -.
DR Genew; HGNC:9385; PRKAG1.
DR MIM; 602742; -.
DR GO; GO:0004691; F:AMP-dependent protein kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
DR InterPro; IPR000644; CBS_domain.
DR SMART; SM00116; CBS; 4.
DR Ffam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 48 96 CBS 1.
FT DOMAIN 123 177 CBS 2.
FT DOMAIN 197 250 CBS 3.
FT DOMAIN 271 323 CBS 4.
SQ SEQUENCE 331 AA; 37579 MW; 0F22B9CA1DBD87AE CRC64;

Query Match 15.7%; Score 367; DB 1; Length 331;
Best Local Similarity 29.1%; Pred. No. 2.5e-20;
Matches 111; Conservative 74; Mismatches 133; Indels 64; Gaps 14;

Qy 69 LSSPSTPGSRMNMVDNENFQRTVLTSDGTVSEGLTAVSEAAIQISRCRVSEYLNLTQY 128
Db 5 ISSDSPA-----VENEHQETPE-SNSV-----YTSFKSHRCY 39
Qy 129 DLLPDGKVIADINLPVKOSFILHEQGPVAPLWDSFRGQVGLLSPDLFILREJE 188
Db 40 DLIPTSSKLVFDTSLQVKKAFVALVNGVRAAPLWDSKKQSFVGMUTITDFINLHRY- 98
Qy 189 THGSNLTE-BOLETHHTISAMKEAKRQINGRNDQWRPQOHLVHATPYESURDIAVKLQN 247
Db 99 -YKSALVQIYELEHETWRELYLQ-----DSFKP---LWCISPNASLFDVAVSSLRN 148
Qy 248 GISTVPVIYSSSDGSPQLHLASLGLKLCICRYFPKNSGTNLPILNQVPCSPILGSWV 307
Db 149 KIHRLPVIDPESGN-----LYILTHKRLKFL-KLFITEFPKPEFMSKSLLEELQIGTYA 202
Qy 308 PKIGDLNRLPLMRPNASLSSALNMLVQGVSSIPVDNDSLLDTYSRSDITALAKDK 367
Db 203 -----NIAMVRITTPVYVALGIFVCHRVSAVPALPVWDEKGRVWDIYKSFVDINLAEEK 253
Qy 368 VYTHVRLEDEMTHQALQGLQGDANTPFGFNGQRCQMCIRSDPILKVMERLANPQVRRFVI 427
Db 254 TYNL---DVSVTQALQHSRSH-----YFEG---VLKCYLHETLTITNLRVRAEVRHLVV 302
Qy 428 VRAGSKRVEGIISLSDFIKFL 449
Db 303 VDENDV-VKGIYSLSLQALV 323

RESULT 3
AAKG BOVIN
ID AAKG BOVIN STANDARD; PRT; 330 AA.
AC P58108;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
```

10-OCT-2003 (Rel. 42, Last annotation update)
5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
(AMPKg).
GN PRKAG1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Benkel B., Kollers S., Fries R., Sazarov A., Yoshida E., Davoren J.,
RA Hickey D.,
RT "Characterization of the bovine AMPK gamma-1 gene."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPIASE. THIS IS A REGULATORY SUBUNIT.
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF329081; AAK19307.1; -
DR InterPro; IPR006444; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 48 96 CBS 1.
FT DOMAIN 123 177 CBS 2.
FT DOMAIN 197 250 CBS 3.
FT DOMAIN 271 323 CBS 4.
SQ SEQUENCE 330 AA; 37481 MW; F130ACBEE2BFE89 CRC64;
Query Match 15.6%; Score 365; DB 1; Length 330;
Best Local Similarity 28.8%; Pred. No. 3.5e-20;
Matches 109; Conservative 73; Mismatches 138; Indels 58; Gaps 13;
QY 73 STPGSRMMDVDNENFQRTVTLSDCTVSEGLRVSEMAIQISRCRVSEVLAHLTCYDLIP 132
DB 3 AVPSGDSYPAVENEHLQETPE-SNSV-----YTSFPMKSHRCYDLIP 43
QY 133 DSGKVALDIMPVQKSHFLHLEQIPVAPLWDSFGQFVGLLSPLDFTLILRELETHGS 192
DB 44 TSSKLAVFTSLQVKKAFALVTNGVRAAPLWDSKKQSPVGMLTITDFINILHRY--YKS 101
QY 193 NLTE-EQLTHTISAWKAKQRTNGSDNSQWRPQCHLHATPYESLRDIANKLNGIST 251
DB 102 ALVQYIELEHKIETWREVFLO-----DSFKPZ---LVCISPNASLFDVAVSLIRNKIHR 152
QY 252 VPVIYSSSDSGSFPQLLHSLASGLIKTCICRYFKNSTGNLTNCPVCSIPGSPVVKIG 311
DB 153 LPVIDPESGN-----TVILTHKRLKFL-KLFTTEFPKPEPMKSLBELQIGTYA---- 202
QY 312 DLNSRPLAMLRPNASLSSALNMLVQAGVSSIFIVDNDPSLLDTYGRSDITALAKDKVYTH 371
DB 203 -----NIAMVRTTTPYVALGIFVQHRVSLPVPVDEKGRVVDIYSKFDVINLAEEKTYNN 257
QY 372 VLDEMTHQALQLQGDQANTPEGFNGQRCQCLSDPLLKVMERLAMPGRVVRVIVZAG 431
DB 256 L---DVSVTKALQHRSH-----YFEG---VLKCYLHETLETIINRLVEAEVRLLVVVDEN 306

QY 432 SKRVEGIISLSDIFKFL 449
DB 307 DV-VKGIIVSLSDILQALV 323
RESULT 4
ID AAKG RAT STANDARD; PRT; 330 AA.
AC P80385;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE (AMPKg).
GN PRKAG1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Wistar;
RX MEDLINE=96215327; PubMed=8626596;
RA Woods A., Cheung P.C.F., Smith F.C., Davison M.D., Scott J.,
RA Berri R.K., Carling D.,
RT "Characterization of AMP-activated protein kinase beta and gamma
RT subunits. Assembly of the heterotrimeric complex in vitro."
RL J. Biol. Chem. 271:10282-10290(1996).
RN [2]
RP SEQUENCE OF 8-330 FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96224074; PubMed=8621499;
RA Gao G., Fernandez C.S., Stapleton D., Auster A.S., Widmer J.,
RA Dyck J.R.B., Kemp B.E., Witters L.A.;
RT "Non-catalytic beta- and gamma-subunit isoforms of the
RT 5'-AMP-activated protein kinase."
RL J. Biol. Chem. 271:8675-8681(1996).
RN [3]
RP SEQUENCE OF 48-330 FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95050763; PubMed=7961907;
RA Stapleton D., Gao G., Michel B.J., Widmer J., Mitchell K.,
RA Teh T., House C.M., Witters L.A., Kemp B.E.;
RT "Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are
RT homologs of proteins that interact with yeast Snf1 protein kinase."
RL J. Biol. Chem. 269:29343-29346(1994).
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HORMONE-SENSITIVE LIPIASE AND
CC HYDROXYMETHYLGUTARYL-COA REDUCTASE. APPEARS TO ACT AS A METABOLIC
CC STRESS-SENSING PROTEIN KINASE SWITCHING OFF BIOSYNTHETIC PATHWAYS
CC WHEN CELLULAR ATP LEVELS ARE DEPLETED AND WHEN 5'-AMP RISES IN
CC RESPONSE TO FUEL LIMITATION AND/OR HYPOXIA. THIS IS A REGULATORY
CC SUBUNIT.
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -!- TISSUE SPECIFICITY: Highly expressed in heart and brain, also
CC found in kidney, white adipose tissue, lung and spleen.
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X95578; CAA64831.1; -
DR EMBL; U42413; AAC52580.1; -


```

DR PIR; T10759;
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
DR Fatty acid biosynthesis; Repeat; CBS domain.
FT DOMAIN 47 95 CBS 1.
FT DOMAIN 122 176 CBS 2.
FT DOMAIN 136 249 CBS 3.
FT DOMAIN 270 322 CBS 4.
FT CONFLICT 114 114 E -> Q (IN REF. 3).
FT CONFLICT 201 201 A -> P (IN REF. 3).
SQ SEQUENCE 330 AA; 37386 MW; 36031E526C1F1B97 CRC64;

Query Match 15.2%; Score 355.5; DB 1; Length 330;
Best Local Similarity 30.3%; Pred. No. 1.8e-19;
Matches 100; Conservative 67; Mismatches 124; Indels 39; Gaps 11;

QY 121 YLNLHTCYLLPDSGKVIADINLPVKQSFHILHEQIPVAPLWDSFRGQFVGLLSPLDF 180
Db 31 FMKSHRCYLLIPTSSKLIVFDTSLOVKKAFPAFTVNGVRAAPLWDSKQSFVGLLITDF 90
QY 181 ILILRETHGSLNTE-EOLETHITISAWKEAKQTNGRNDOSQWRPOOHLVHATPYESLRD 239
Db 91 INILHRY--YKSAVQIYELEHETWREVLQ-----DSFKP---LVCISPNASLED 139
QY 240 IAVKLLONGISTVPVIYSSSDGSPQLHLASLSGLKXICICFYKXNSICNLPLNQPV 299
Db 140 AVSSLRNKIHLRPVIDPSGN-----TLYLTHKRLKFL-KLFTIEFPKPFMSKSL 193
QY 300 SIPLGSWEPEKIGDLSRPLMLRPNASLSALNMLVQAGVSSIPVDDNDSLDITYSRSD 359
Db 194 ELIQIYIA-----NIAWRTTTPVYVALGIFVQHRVSAIPVDEKGRVVDIYSKED 244
QY 360 ITALAKDKVYTHVRLDEMTIHOALQIQDANTPFQFNGRCQMCILRSDPLLKMERLAN 419
Db 245 VINLAAEKTNNL--DVSVTKALQHRSH-----YFEG--VLKCYLHETLETIINRLVE 293
QY 420 PGVRRVFIIVEAGSKRVEGIISLSDIFKELL 449
Db 294 AEVHRLVVVDEHDV-VKGIIVSLSDILQALV 322

RESULT 5
AAKG MOUSE
ID AAKG MOUSE STANDARD; PRT; 330 AA.
AC 054950;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain)
DE (AMPKg).
GN PRKAG1 OR PKAAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21203559; PubMed=11306812;
RA Shamsadin R., Jantsan K., Adham I., Engel W.;
RT "cloning, organisation, chromosomal localization and expression
RT analysis of the mouse Prkag1 gene";
RL Cytogenet. Cell Genet. 92:134-138 (2001).
CC -1- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT.
CC -1- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -1- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.

```

QY	296	QVCSITPLGSWPVKIGDIAMSPLMLRPKVASLSALNMLVOAGVSSIPVDDNDSLDITY	350
DB	321	RTIQDIGITGF-----RDIAVDFATPIALDIFVDRVSALPVNVECGCVGLY	371
QY	356	SRSDDITALAKDYIVHVRLEMTIHQALQIQDANTFFGFFNQRCQRCMLRS-----D	408
DB	372	SFPDVIHLAAQTYNHL---DMSVGEALR-----QR-TLCLEGVLSQCPHE	413
QY	409	PLIKWNERLANGVRRVVFIVEAGSKRVEGIISLSDIFKELL	449
DB	414	SLGEVIDIRAREQVHRLVLVDE-TQHLGTVLSLSDILQALV	453

RESULT 7

AAKL_PIG

AD

AAKL_PIG

IC

Q9MYP4;

16-OCT-2001

(Rel. 40, Created)

DT

10-OCT-2003

(Rel. 42, Last sequence update)

DT

15-MAR-2004

(Rel. 43, Last annotation update)

DE

5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain)

DE

(AMPK gamma3).

GN

PRKAG3.

OS

Sus scrofa (Pig).

OC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX

NCBI_TaxID=9823;

RN

[1]

RP

SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT RN(-) GLN-250.

RC

TISSUE=Skeletal muscle;

RC

MEDLINE=20280150; PubMed=10818001;

RA

Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Thelander M.,

RA

Rogel-Gaillard C., Paul S., Iannucelli N., Rask L., Ronne H.,

RA

Lundstroem K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P.,

RA

Andersson L.;

RT

*A mutation in PRKAG3 associated with excess glycogen content in pig

RT

skeletal muscle.";

RT

Science 288:1248-1251(2000).

RN

[2]

RP

SEQUENCES FROM N.A. (ISOFORM 2).

RC

TISSUE=Skeletal muscle;

RA

Milan D., Jeon J.-T., Looft C., Amarger V., Robic A., Reinsch N.,

RA

Rogel-Gaillard C., Paul S., Gellin J., Lundstrom K., Reinsch N.,

RA

Kalm E., Le Roy P., Chardon P., Andersson L.;

RA

Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.

CC

-I- FUNCTION: AMPK is responsible for the regulation of fatty acid

CC

synthesis by phosphorylation of acetyl-CoA carboxylase. Also

CC

regulates cholesterol synthesis via phosphorylation and

CC

inactivation of hydroxymethylglutaryl-CoA reductase and hormone-

CC

sensitive lipase. This is a regulatory subunit. It may play a role

CC

in the regulation of energy metabolism in skeletal muscle.

CC

-I- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a

CC

gamma non-catalytic regulatory subunits.

CC

-I- ALTERNATIVE PRODUCTS:

CC

Event-Alternative splicing; Named isoforms=2;

CC

Name=2;

CC

Isoidc=Q9MYP4-1; Sequence=Displayed;

CC

Name=1;

CC

Isoidc=Q9MYP4-2; Sequence=VSP_008059;

CC

-I- TISSUE SPECIFICITY: Muscle.

CC

-I- DISEASE: Defects in PRKAG3 re the cause of the RN- phenotype which

CC

is associated with excess glycogen content (about 70%) in skeletal

CC

muscle. This mutation originated in the hamshire breed pigs and

CC

has beneficial effects on meat content but detrimental effects on

CC

processing yield. Thus, this mutation is of considerable economic

CC

significance in the pig breeding industry.

CC

-I- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma

CC

subunit family.

CC

-I- SIMILARITY: Contains 4 CBS domains.

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC

between the Swiss Institute of Bioinformatics and the EMBL Outstation

CC

at the European Bioinformatics Institute. There are no restrictions on its

CC

use.

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/send_email_to_license@isb-sib.ch).

EMBL; AF214520; AAF73988.2; -	
EMBL; AF214521; AAF73989.1; -	
InterPro; IPR000644; CBS_domain	
PFam; PF00571; CBS; 4	
SMART; SM01116; CBS; 4	
Fatty acid biosynthesis; Repeat; CBS domain; Disease mutation;	
Alternative splicing.	
DOMAIN 222 276	CBS 1.
DOMAIN 303 357	CBS 2.
DOMAIN 378 431	CBS 3.
DOMAIN 450 503	CBS 4.
VARSPPLIC 1 50	Missing (in isoform 1).
	/FTD=VSP_008059.
VARIANT 250 250	R -> Q (IN RN-).
SEQUENCE 514 AA; 56789 MW; 8CE025FBF93E4AE CRC64;	

RESULT 8

SNF4_YEAST	STANDARD;	PRT;	322 AA.
ID	SNF4_YEAST		
AC	P12904;		
DT	01-OCT-1989 (Rel. 12, Created)		
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Nuclear protein SNF4 (Regulatory protein CAT3).		
DE	SNF4 OR CAT3 OR YGL115W.		
GN	Oosaccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes.		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
OX	NCBI_TaxID=4932;		
RN	[1]		
RX	SEQUENCE FROM N.A.		
RP	MEDLINE=89006284; PubMed=3049255;		
RX	Schuessler H.-J., Entian K.-D.;		
RT	"Molecular characterization of yeast regulatory gene CAT3 necessary		
RT	for glucose derepression and nuclear localization of its product.;"		
RL	Gene 67:247-257(1988).		
RL	[2]		
RN	SEQUENCE FROM N.A.		
RP	MEDLINE=90097921; PubMed=2481228;		
RX			

InterPro:	IPR000644;	CBS_domain.
DR	Pfam; PF00571; <td>CBS; 4.</td>	CBS; 4.
DR	SMART; SM00116; <td>CBS; 4.</td>	CBS; 4.
KW	Carbohydrate metabolism; Transcription regulation; Nuclear protein;	
KW	Repeat; CBS domain.	
FT	DOMAIN	35 89 CBS 1.
FT	DOMAIN	117 175 CBS 2.
FT	DOMAIN	192 246 CBS 3.
FT	DOMAIN	259 318 CBS 4.
SQ	SEQUENCE	322 AA; 36401 MW; 518387E346EE9561 CRC64;
Query Match	13.0%;	Score 303.5; DB 1; Length 322;
Best Local Similarity	27.9%;	Pred. No. 1.4e-15;
Matches	95; Conservative	64; Mismatches 129; Indels 53; Gaps 12;
Qy	118	VSEYLNLCYTDLLPDSGKVIADINLPVKQSFHLHREQIG?VAPLWDSFRGQFVGLLSP 177
Db	22	IRKFLNSK?SYDVLPSVYELIVLDTSLLVKSLNVLQNSIVSAPLWDSK?SRFAGLLTT 81
Qy	178	LQFILRLRELTGNSNLTEQLETHTI SAWKEAKQ?TNGRNDQWR-----PQCHLVHATPY 234
Db	82	TOPINV---IOVYFNSPDKFLVD-----KQLDGLKDIERALGVDLDTASIHPS 129

```
QY 235 ESURDRAVXKLQNGISTVPVIVSSSDGSPFOLLHLASLGILKCI--CR---YFKNST 288
Db 130 RPLFEACLMLESRSGRIPLI-DQDETHREIVSVLTQYRIILKFPVALNCRETHFLKIPI 188
QY 289 GNLPIILQPCVCSIPLGSWPKIGDLNSRPLAMLRPNASLSALNMLVOAGVSSIPVDDN 348
Db 189 GELNITIQD-----NMKSCQMTTP---VIDVQMLTQGRVSSVPIIDEN 229
QY 349 LSLIDTYSRSDITALAKDKVYTVRLDEMTHQALQIQDANTPFGFNGRCOMCLUSD 408
Db 230 GYLINVEAYDVILGKGIYNDLSL---SVGZALMRSDD-----FEG--VYTCKND 278
QY 409 ELKVMERLANPGVRRVFIYVAGSKRVEGIIISIDIEFKLL 449
Db 279 KLSTINDNIRKARVHRFFVVD-DVGRLVGVLTLSDLIKYL 318

RESULT 9
YL28-SCHPO STANDARD; PRT; 334 AA.
ID YL28 SCHPO
AC Q10343; Q9UTJ1;
DT 01-OCT-1996 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein GL556.08c in chromosome 1.
GN SPAC1556.08C OR SPAC1F12.01C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX KCEI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown J., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton C., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward T., Volkart G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Welter J., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Reinhardt R., Pohl T.M.,
RA Borzym K., Langer J., Beck A., Jehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purcell B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shipkovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; Z69944; CAA93805.1; --
DR EMBL; AL132984; CAB61219.1; --
DR GenBank Sfam; SPAC1556.08c; --
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Hypothetical protein; Repeat; CBS domain.
FT DOMAIN 33 81 CBS 1.
FT DOMAIN 118 170 CBS 2.
FT DOMAIN 194 241 CBS 3.
FT DOMAIN 266 314 CBS 4.
SQ SEQUENCE 334 AA; 37427 MW; 8D32CC6CE53F7916 CRC64;

Query Match 12.3%; Score 289; DB 1; Length 334;
Best Local Similarity 26.0%; Pred. No. 1.9e-14; Indels 36; Gaps 12;
Matches 87; Conservative 74; Mismatches 137;

QY 118 VSEYLNLHTICVDLLPDSGKVIADINLFPVQSFILHEQGIPEVAPLWDSFPGQFVGLLSP 177
Db 15 IQAFIRSETSYDVLPTSFRLIVFDVTLFVKTSLSLLTLNNIVSAPLWDSFPGQFVGLLSP 74
QY 178 LDFILIRELETHGSLNTE--EQLETHTISHAKKAKRQNGRNDNSQWRPQOHLVHATPYE 235
Db 75 ADFVNVIK-YYIQSSSFPEALAEIDKFRLLGLREVERKIGA-----IPPETIYYVH--PMH 126
QY 236 SLRDIYVKKLQNGISTVPVIVSSSDGSPFOLLHLASLGILKCIYFKNSTGMLPLN 295
Db 127 SLMDACLAMSKSRARRIPLIDVDGETGS-EMIVSVLTQYRIILKFIISMCKETA----MLR 181
QY 296 QPVCSIPLGSWPKIGDLNSRPLAMLRPNASLSALNMLVOAGVSSIPVDDNDSLLDTY 355
Db 182 VPLNQMTIGTW-----SNLATASMETKVDYVIMKLAENKISAVEIVNSEGLINNVY 232
QY 356 SRSDITALAKDKVYTVRLDEMTHQALQIQDANTPFGFNGRCOMCLASDILLKYWE 415
Db 233 ESDVHMLIQDGYSNL---DLSVGEAL-LKRPAN-----FDG--VHTCRATDRLDGIFD 281
QY 416 RLAMPGVRRVFIYVAGSKRVEGIIISIDIEFKLL 449
Db 282 AIKHSRVHRLFPVVDENLK-LEGILSLADILNYII 314

RESULT 10
SNF4_KJUL4 STANDARD; PRT; 328 AA.
ID SNF4_KJUL4
AC Q9B869;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nuclear protein SNF4.
GN SNF4.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RA Tomasini L., Ferrero I., Goffrini P.;
RT "Molecular characterization of KLSNF4 gene."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: This protein causes expression of glucose-repressible
CC genes upon glucose deprivation. It interacts and has functional
CC relationship to the protein-kinase SNF1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase gamma
CC subunit family.
CC -!- SIMILARITY: Contains 4 CBS domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```


DT CC the European Bioinformatics Institute. There are no restrictions on its
DT CC use by non-profit institutions as long as its content is in no way
DT CC modified and this statement is not removed. Usage by and for commercial
DT CC entities requires a license agreement (see <http://www.isb-sib.ch/announcements>
DT CC or send an email to license@isb-sib.ch).
DE CC -----
DE GN EMBL; U42411; RAC52579.1; --
OS Rattus norvegicus (Rat).
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RY [1]
RP SEQUENCE FROM N.A.; AND PARTIAL SEQUENCE.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96224074; PubMed=8622499;
RA Gao G., Fernandez C.S., Stapleton D., Austen A.S., Widmer J.,
RA Dyck J.R.B., Kemp B.E., Witters L.A.;
RA "Non-catalytic beta- and gamma-subunit isoforms of the
RT 5'-AMP-activated protein kinase";
RL J. Biol. Chem. 271:8675-8681(1996).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Wistar;
RX MEDLINE=96215327; PubMed=8626596;
RA Woods A., Cheung P.C.F., Smith F.C., Davison M.D., Scott J.,
RA Beri R.K., Carling D.;
RA "Characterization of AMP-activated protein kinase beta and gamma
RT subunits. Assembly of the heterotrimeric complex in vitro";
RL J. Biol. Chem. 271:10282-10290(1996).
RN [3]
RN SEQUENCE OF 35-158 FROM N.A.; AND SEQUENCE OF 35-71; 78-82 AND 89-158.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95050763; PubMed=7961907;
RA Stapleton D., Gao G., Mitchell B.J., Widmer J., Mitchell K.I.,
RA Teh T., House C.M., Witters L.A., Kemp B.E.;
RA "Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are
RT homologs of proteins that interact with yeast Snf1 protein kinase";
RL J. Biol. Chem. 269:29343-29346(1994).
RN [4]
RN PARTIAL SEQUENCE, MYRISTOYLATION, AND PHOSPHORYLATION OF SER-23;
RP SER-24; SER-107 AND SER-181.
RX MEDLINE=97450976; PubMed=9305909;
RA Mitchell K.I., Mitchell B.J., House C.M., Stapleton D., Dyck J.,
RA Gamble J., Ullrich C., Witters L.A., Kemp B.E.;
RA "Posttranslational modifications of the 5'-AMP-activated protein
RT kinase beta subunit";
RL J. Biol. Chem. 272:24475-24479(1997).
RN [5]
RN MUTAGENESIS, MYRISTOYLATION, AND PHOSPHORYLATION OF SER-23;
RP SER-24; SER-107 AND SER-181.
RX MEDLINE=21092935; PubMed=1171104;
RA Warden S.M., Richardson C., O'Donnell J. Jr., Stapleton D., Kemp B.E.,
RA Witters L.A.;
RA "Post-translational modifications of the beta-1 subunit of
RT AMP-activated protein kinase affect enzyme activity and cellular
RT localization";
RL Biochem. J. 354:275-283(2001).
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID
CC SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO
CC REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND
CC INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-
CC SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT. MAY BE A POSITIVE
CC REGULATOR OF AMPK ACTIVITY. IT MAY ALSO SERVE AS AN ADAPTER
CC MOLECULE FOR THE CATALYTIC ALPHA-SUBUNIT.
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney, heart, white
CC adipose tissue, lung and spleen.
CC -!- PTM: Phosphorylated when associated with the catalytic subunit.
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase beta
CC subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC -----

CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see <http://www.isb-sib.ch/announcements>
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; U42411; RAC52579.1; --
CC DR EMBL; X95577; CAA64830.1; --
CC DR InterPro; IPR006828; AMPKBI.
CC DR Pfam; PF04739; AMPKBI; 1.
CC KW Fatty acid biosynthesis; Phosphorylation; Myristate; Lipoprotein.
CC FT INIT MET 0 0
CC FT LIPID 1 1 N-myristoyl glycine.
CC FT MOD RES 23 23 PHOSPHORYLATION (AUTO-).
CC FT MOD RES 24 24 PHOSPHORYLATION (AUTO-).
CC FT MOD RES 107 107 PHOSPHORYLATION (AUTO-).
CC FT MOD RES 181 181 PHOSPHORYLATION.
CC FT MOD RES 181 181 PHOSPHORYLATION.
CC FT CONFLICT 25 25 G -> E (IN REF. 1).
CC FT CONFLICT 51 51 M -> I (IN REF. 3; AA SEQUENCE).
CC SQ SEQUENCE 269 AA; 30263 MW; 6245087E57E581E1 CRC64;
Query Match 5.3%; Score 124; DB 1; Length 269;
Best Local Similarity 32.5%; Pred. No. 0.036;
Matches 27; Conservative 14; Mismatches 26; Indels 16; Gaps 2;
QY 9 PIEGCPVFOAICSLSPGIIHEVYKFFVDGEWRHDSRQPTISGEFGIVNTLYL----- 59
DB 103 PLTRSQNNFVALDLPRGEHOYKFFVDGQWTHDPSEPIVTSQLGIVNNIIQVKKTDFFVF 162
QY 60 -----TREYNQINTLSPSTPG 76
DB 163 DALMVDQSQCSDVSELSS-SPPG 184
RESULT 13
ID AAKX HUMAN STANDARD; PRT; 269 AA.
AC Q9Y478; Q9Y478; Q9Y478; Q9Y478; Q9Y478;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, beta-1 subunit (AMPK beta-1 chain)
DE (AMPKb).
GN PRKAB1 OR AMPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Carling D.;
RT "Non-catalytic beta and gamma subunits isoforms of the AMP-activated
RT protein kinase";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97367941; PubMed=9224708;
RA Stapleton D., Woolfitt E., Mitchell K., Nicholl J.K.,
RA Fernandez C.S., Mitchell B.J., Witters L.A., Power D.A.,
RA Sutherland G.R., Kemp B.E.;
RA "AMP-activated protein kinase isoenzyme family: subunit structure and
RT chromosomal location";
RL FEBS Lett. 409:452-456(1997).
RN [3]
RN SEQUENCE FROM N.A.
RA Yamagata K., Oda N., Furuta H., Vaxillaire M., Southam L., Boriraj V.,
RA Chen X., Oda Y., Takeda J., Yamada S., Nishigori H., Lebeau M.M.,
RA Lathrop M., Cox R.D., Bell G.I.;
RA "Transcription map of the 5CM region surrounding the hepatocyte
RT nuclear factor-1a/MODY3 gene on chromosome 12";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]

RP SEQUENCE FROM N.A.
RA Wang X., Yu L., Tu Q.;
RT "Cloning and expression of the complete mRNA coding human AMP-
RL activated protein kinase."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Connell M., Goela D., Harper M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA TISSUE=Lung, and Muscle;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Jasin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska J., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: AMPK is responsible for the regulation of fatty acid
CC synthesis by phosphorylation of acetyl-CoA carboxylase. Also
CC regulates cholesterol synthesis via phosphorylation and
CC inactivation of hydroxymethylglutaryl-CoA reductase and hormone-
CC sensitive lipase. This is a regulatory subunit, may be a positive
CC regulator of AMPK activity. It may also serve as an adaptor
CC molecule for the catalytic alpha-subunit.
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -!- PTM: Phosphorylated (By similarity).
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase beta
CC subunit family.
CC -!- CAUTION: Ref.4 sequence differs from that shown due to a
CC frameshift in position 244.
CC -!- CAUTION: Ref.5 sequence differs from that shown due to erroneous
CC gene model prediction.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; AJ224515; CAA12024.1; -;
CC ENBL; Y12556; CAA73146.1; -;
CC ENBL; U83994; AAD09237.1; -;
CC ENBL; U87276; AAD0625.1; -;
CC ENBL; U87271; AAD0625.1; JOINED.
CC ENBL; U87272; AAD0625.1; JOINED.
CC ENBL; U87273; AAD0625.1; JOINED.
CC ENBL; U87274; AAD0625.1; JOINED.
CC ENBL; U87275; AAD0625.1; JOINED.
CC ENBL; AF022116; AAC98897.1; ALT FRAMES.
CC ENBL; AC002563; AAB71326.1; ALT_SEQ.
CC ENBL; BC001007; AAH01007.1; -;
CC ENBL; BC001056; AAH01056.1; -;
CC ENBL; BC001823; AAH01823.1; -;
CC ENBL; BC017671; AAH17671.1; -;

DR PIR; T09514; T09514.
DR Genew; HGNC:9378; PRKAB1.
DR MIM; 602740; -;
DR GO; GO:0007185; P:signal transduction; TAS.
DR InterPro; IPR006828; AMPKBI.
DR Pfam; PF04739; AMPKBI; 1.
KW Fatty acid biosynthesis; Phosphorylation; Myristate; Lipoprotein.
FT INIT MET 0 0
FT LIPID 1 1 N-myristoyl glycine (By similarity).
FT MOD RES 23 23 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD RES 24 24 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD RES 107 107 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD RES 181 181 PHOSPHORYLATION (BY SIMILARITY).
FT CONFLICT 9 9 A -> G (IN REF. 2 AND 4).
FT CONFLICT 14 14 G -> A (IN REF. 1).
FT CONFLICT 19 19 P -> A (IN REF. 2 AND 4).
FT CONFLICT 21 21 R -> K (IN REF. 3).
FT CONFLICT 55 55 E -> Y (IN REF. 3).
SQ SEQUENCE 269 AA; 30251 MW; F08BCF39D5AAF861 CRC64;
Query Match 5.3%; Score 123; DB 1; Length 269;
Best Local Similarity 34.7%; Pred. No. 0.043; 20; Indels 16; Gaps 2;
Matches 26; Conservative 13; Mismatches 13;
Oy 17 FQACISLPGIHEHYKFFVDGSEWRHDERQPTISGEFGIVNTLYL-----TR 61
Db 111 FVALDLPEGEHQYKFFVDGQWTHDPSEPIVTSQLTGNNIIQVKKTFEVDALMVDSQ 170
Oy 62 EYNQINTLSSESTPG 76
Db 171 KCDVSELS-SPFG 184
RESULT 14
AKAC RAT
ID AKAC RAT STANDARD; PRT; 271 AA.
AC Q9QZH4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5'-AMP-activated protein kinase, beta-2 subunit (AMPK beta-2 chain).
GN PRKAB2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Skeletal muscle;
RX MEDLINE=20012956; PubMed=10544261;
RA Chen Z., Heierhorst J., Mann R.J., Mitchell K.I., Michell B.J.,
RA Witters L.A., Lynch G.S., Kemp B.E., Stapleton D.;
RT "Expression of the AMP-activated protein kinase beta1 and beta2
RT subunits in skeletal muscle."
RL FEBS Lett. 460:343-348 (1999).
CC -!- FUNCTION: AMPK is responsible for the regulation of fatty acid
CC synthesis by phosphorylation of acetyl-CoA carboxylase. Also
CC regulates cholesterol synthesis via phosphorylation and
CC inactivation of hydroxymethylglutaryl-CoA reductase and hormone-
CC sensitive lipase. This is a regulatory subunit, may be a positive
CC regulator of AMPK activity. It may also serve as an adaptor
CC molecule for the catalytic alpha-subunit.
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC gamma non-catalytic regulatory subunits.
CC -!- PTM: Phosphorylated when associated with the catalytic subunit.
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase beta
CC subunit family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC  -|- FUNCTION: AMPK is responsible for the regulation of fatty acid
CC  synthesis by phosphorylation of acetyl-CoA carboxylase. Also
CC  regulates cholesterol synthesis via phosphorylation and
CC  inactivation of hydroxymethylglutaryl-CoA reductase and hormone-
CC  sensitive lipase. This is a regulatory subunit, may be a positive
CC  regulator of AMPK activity. It may also serve as an adaptor
CC  molecule for the catalytic alpha-subunit.
CC  -|- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a
CC  gamma non-catalytic regulatory subunits.
CC  -|- PTM: Phosphorylated when associated with the catalytic subunit.
CC  -|- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase beta
CC  subunit family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AJ224538; CAA12030.1; -
DR  EMBL; AF504543; AAM74153.1; -
DR  EMBL; AF504538; AAM74153.1; JOINED.
DR  EMBL; AF504539; AAM74153.1; JOINED.
DR  EMBL; AF504540; AAM74153.1; JOINED.
DR  EMBL; AF504541; AAM74153.1; JOINED.
DR  EMBL; AF504542; AAM74153.1; JOINED.
DR  EMBL; BC053610; AAH53610.1; -
DR  Genew; HGNC:9379; PRKAB2.
DR  MIM; 602741; -
DR  GO; GO:0007165; P:signal transduction; TAS.
DR  InterPro; IPR006828; AMPKBE1.
DR  Pfam; PF04739; AMPKBE1; 1
DR  Fatty acid biosynthesis. Phosphorylation.
KW  MOD_RES 183 183 PHOSPHORYLATION (PARTIAL) (BY
FT  SIMILARITY).
SQ  SEQUENCE 272 AA; 30302 MW; 42B23BD70B92519C CRC64;

Query Match 5.2%; Score 121; DB 1; Length 272;
Best Local Similarity 48.8%; Pred. No. 0.561;
Matches 20; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Qy 17 FOAICSLSPGIEHKYKFFVDCGEWRHDEROPTISGEFGVNTLL 57
Db 112 FVAILLDPGEHQKFFVDGQWHDPEPVTQSGLTNNL 152

Search completed: July 7, 2004, 17:51:20
Job time : 19 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:49:44 ; Search time 52 Seconds

(without alignments)

2699.788 Million cell updates/sec

Title: US-09-857-525C-2

Perfect score: 2342

Sequence: 1 TREHLPMSPIECPTVFQAI.....SKRVEGISLSIDIEKFLLSL 451

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/USC6_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/USC6_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	95.3	514	12	US-10-425-114-38422 Sequence 38422, A
2	2230	95.2	514	12	US-10-425-114-64906 Sequence 64906, A
3	2178	93.0	423	12	US-10-425-114-38456 Sequence 38456, A
4	1944	83.0	723	16	US-10-437-963-105356 Sequence 105356, A
5	1802	76.9	365	12	US-10-425-114-69679 Sequence 69679, A
6	1527.5	65.2	493	12	US-10-424-599-187700 Sequence 187700, A
7	1492	63.7	302	12	US-10-425-114-58697 Sequence 58697, A
8	1487	63.5	461	12	US-10-424-599-225531 Sequence 225531, A
9	1325	56.6	368	12	US-10-425-114-54796 Sequence 54796, A
10	923.5	39.4	263	12	US-10-424-599-142943 Sequence 142943, A
11	920	39.3	299	12	US-10-424-599-142947 Sequence 142947, A
12	908.5	38.8	2248	16	US-10-437-963-193213 Sequence 193213, A
13	820.5	35.0	451	16	US-10-437-963-168583 Sequence 168583, A
14	778	33.2	166	16	US-10-437-963-125316 Sequence 125316, A
15	761	32.5	477	12	US-10-425-114-50202 Sequence 50202, A

16	741.5	31.7	230	12	US-10-424-599-142942	Sequence 142942, A
17	580.5	24.8	232	12	US-10-425-114-38434	Sequence 38434, A
18	425.5	18.2	255	12	US-10-425-114-57823	Sequence 57823, A
19	373	15.9	328	10	US-09-769-970-7	Sequence 7, Appli
20	367	15.7	331	10	US-09-769-970-21	Sequence 21, Appl
21	359	15.3	344	9	US-09-925-297-461	Sequence 461, App
22	359	15.3	1207	13	US-10-108-605-71	Sequence 187709, A
23	356	15.2	193	12	US-10-424-599-187709	Sequence 187709, A
24	340	14.5	489	9	US-09-826-581-6	Sequence 6, Appli
25	340	14.5	489	16	US-10-473-670-15	Sequence 15, Appl
26	340	14.5	489	16	US-10-705-137-6	Sequence 6, Appli
27	335.5	14.3	464	12	US-09-950-022-8	Sequence 8, Appli
28	334.5	14.3	464	12	US-09-950-022-10	Sequence 10, Appl
29	334.5	14.3	464	12	US-09-950-022-2	Sequence 2, Appli
30	334.5	14.3	464	12	US-09-950-022-4	Sequence 4, Appli
31	334.5	14.3	464	12	US-09-950-022-6	Sequence 6, Appli
32	297.5	12.7	120	12	US-10-424-599-148301	Sequence 148301, A
33	231	9.9	136	12	US-10-424-599-187699	Sequence 187699, A
34	173	7.4	149	15	US-10-264-237-1653	Sequence 1653, Ap
35	153	6.5	92	12	US-10-424-599-197913	Sequence 197913, A
36	151	6.4	57	12	US-10-424-599-210108	Sequence 210108, A
37	133	5.7	83	12	US-10-424-599-187707	Sequence 187707, A
38	130.5	5.6	180	14	US-10-106-698-6572	Sequence 6572, Ap
39	115.5	4.9	443	12	US-10-424-599-223551	Sequence 223551, A
40	111.5	4.8	181	12	US-10-276-774-1611	Sequence 1611, Ap
41	110.5	4.7	446	12	US-10-282-122A-69591	Sequence 69591, A
42	108.5	4.6	425	16	US-10-437-963-161228	Sequence 161228, A
43	107.5	4.6	4910	12	US-10-346-863-21	Sequence 21, Appl
44	107	4.6	476	12	US-10-282-122A-60388	Sequence 60388, A
45	106.5	4.5	1274	16	US-10-437-963-196165	Sequence 196165, A

ALIGNMENTS

RESULT 1

US-10-425-114-38422
; Sequence 38422, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Lib, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38422
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700172325_FLI.pep
US-10-425-114-38422

Query Match	95.3%	Score 2231;	DB 12;	Length 514;
Best Local Similarity	95.5%;	Pred. No. 4.8e-213;		
Matches 429;	Conservative 8;	Mismatches 12;	Indels 0;	Gaps 0;
Qy	3	EHLPMSPTEGCTVFQAI	CSLSGCIHEYKPFVDCGWHDERQPTISGEFGWNTLYLTRE	62
Db	66	EHLPMSPTEGCTVFQAI	CSLSGCIHEYKPFVDCGWHDERQPTISGEFGWNTLYLTRE	125
Qy	63	YNQINTLSSPTGSRMNMVDNENFQRTVLTSDGTVSEGLRVSEAAIQISRCRVSEYL		122
Db	126	FNQINALLNPTGSRMNMVDNENFQHTVLTSDGTPEGIVRVSEAAIQISRCRVSEYL		185
Qy	123	NLHTCYDLLPDSKGVIALDINLPVQSFHILHEGQIPVAPLWDSFRGQVGLLSPLDIL		182

Db 186 NLHTCYDLPDSKVALDINLPVKQSFHILHEQGPVAPLWDSFRGQFVGLLSPLDFIL 245
Qy 183 ILRELETHGSLNTEQLETHITISAWKEAKQTNGRNDQWRPQOHLVHATPYESLRDIIV 242
Db 246 ILRELETHGSLNTEQLETHITISAWKEAKQTNGRNDQWRPQOHLVHATPYESLRDIIV 305
Qy 243 KLLONGISTVPVIYSSSSDGSFPQLHLASLSGLIKICRYFNKSTGNLPILOPVCISIP 302
Db 306 KLLANDISTVPVIYSSSSDGSFPQLHLASLSGLIKICRYFNKSTGNLPILOPVCISIP 365
Qy 303 LGSWPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNDSLDTSRSDITA 362
Db 366 LGSWPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNDSLDTSRSDITA 425
Qy 363 LAKDKVYTHVRLDEMTIHQALQIQDANTPFGFNQRCQMLRSDPPLLKVMERLANPGV 422
Db 426 LAKDKVYTHVRLDEMTIHQALQIQDANTPFGFNQRCQMLRSDPPLLKVMERLANPGV 485
Qy 423 RRVFIVEAGSKRVEGIISLSDIFKFLLSL 451
Db 486 RRVFIVEAGSKRVEGIISLSDIFKFLLSL 514

RESULT 2

US-10-425-114-64906
; Sequence 64906, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73:28
; SEQ ID NO 64906
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4729-C18-A12_FLI.pep
US-10-425-114-64906

Query Match 95.2%; Score 2230; DB 12; Length 514;
Best Local Similarity 95.3%; Pred. No. 6e-213;
Matches 428; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

Qy 3 EHLPMSPIEGCTVFQAICSLSPGHEHYKFPVDEWRHDERQPTISGEFIVNTLYLTRE 62
Db 66 EHLPMSPVEGCTVFQAICSLSPGHEHYKFPVDEWRHDERQPTISGEFIVNTLYLTRE 125
Qy 63 YNQINTLSSPSTGSRNMVDNENFQRTVTLSDGTVSEGTILRSEAAIQISRCRVSEYL 122
Db 126 FNOINALLNFTSGSRNMVDNENFQRTVTLSDGTIPBGTIVRSEAAIQISRCRVSEYL 185
Qy 123 NLHTCYDLPDSKVALDINLPVKQSFHILHEQGPVAPLWDSFRGQFVGLLSPLDFIL 182
Db 186 NLHTCYDLPDSKVALDINLPVKQSFHILHEQGPVAPLWDSFRGQFVGLLSPLDFIL 245
Qy 183 ILRELETHGSLNTEQLETHITISAWKEAKQTNGRNDQWRPQOHLVHATPYESLRDIIV 242
Db 246 ILRELETHGSLNTEQLETHITISAWKEAKQTNGRNDQWRPQOHLVHATPYESLRDIIV 305
Qy 243 KLLONGISTVPVIYSSSSDGSFPQLHLASLSGLIKICRYFNKSTGNLPILOPVCISIP 302
Db 306 KLLANDISTVPVIYSSSSDGSFPQLHLASLSGLIKICRYFNKSTGNLPILOPVCISIP 365

Qy 303 LGSWPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNDSLDTSRSDITA 362
Db 366 LGSWPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNDSLDTSRSDITA 425
Qy 363 LAKDKVYTHVRLDEMTIHQALQIQDANTPFGFNQRCQMLRSDPPLLKVMERLANPGV 422
Db 426 LAKDKVYTHVRLDEMTIHQALQIQDANTPFGFNQRCQMLRSDPPLLKVMERLANPGV 485
Qy 423 RRVFIVEAGSKRVEGIISLSDIFKFLLSL 451
Db 486 RRVFIVEAGSKRVEGIISLSDIFKFLLSL 514

RESULT 3

US-10-425-114-38456
; Sequence 38456, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73:28
; SEQ ID NO 38456
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700342215_FLI.pep
US-10-425-114-38456

Query Match 93.0%; Score 2178; DB 12; Length 423;
Best Local Similarity 99.5%; Pred. No. 6.7e-208;
Matches 421; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 29 EYKFFVDGWRHDERQPTISGEFIVNTLYLTREYNQINTLSSPSTGSRNMVDNENF 88
Db 1 EYKFFVDGWRHDERQPTISGEFIVNTLYLTREYNQINTLSSPSTGSRNMVDNENF 60
Qy 89 QRTVTLSDGTVSEGTILRSEAAIQISRCRVSEYMLHTCYDLPDSKVALDINLPVKQ 148
Db 61 QRTVTLSDGTVSEGTILRSEAAIQISRCRVSEYMLHTCYDLPDSKVALDINLPVKQ 120
Qy 149 SFHILHEQGPVAPLWDSFRGQFVGLLSPLDFILILRELETHGSLNTEQLETHITISAWK 208
Db 121 SFHILHEQGPVAPLWDSFRGQFVGLLSPLDFILILRELETHGSLNTEQLETHITISAWK 180
Qy 209 EAKQRTNGRNDQWRPQOHLVHATPYESLRDIIVKLLONGISTVPVIYSSSSDGSFPOLL 268
Db 181 EAKQRTNGRNDQWRPQOHLVHATPYESLRDIIVKLLONGISTVPVIYSSSSDGSFPOLL 240
Qy 269 HLASLSGLIKICRYFNKSTGNLPILOPVCISIPGSRVPMKIGDLSRPLAMLRPNASLS 328
Db 241 HLASLSGLIKICRYFNKSTGNLPILOPVCISIPGSRVPMKIGDLSRPLAMLRPNASLS 300
Qy 329 SALNMLVQAGVSSIPVDDNDSLDTSRSDITALAKDKVYTHVRLDEMTIHQALQIQD 388
Db 301 SALNMLVQAGVSSIPVDDNDSLDTSRSDITALAKDKVYTHVRLDEMTIHQALQIQD 360
Qy 389 ANTFFGFFNGRCQMLRSDPPLLKVMERLANPGVRRVFRVIVEAGSKRVEGIISLSDIFKFL 448
Db 361 ANTFFGFFNGRCQMLRSDPPLLKVMERLANPGVRRVFRVIVEAGSKRVEGIISLSDIFKFL 420
Qy 449 LSL 451
Db 421 LSL 423

RESULT 4

US-10-437-963-105356
; Sequence 105356, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)/B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 105356
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_102608C.1.pep
US-10-437-963-105356

Query Match 83.0%; Score 1944; DB 16; Length 723;

Best Local Similarity 86.4%; Pred. No. 3.4e-184;

Matches 361; Conservative 34; Mismatches 23; Indels 0; Gaps 0;

QY	3	EHLPMSEICPTVFQAICSLSPGIBHYKFPVDEWHRDERTISBFGVINTLYLTRE	62
DB	46	EHLPMSEICPTVFQAICSLSPGIBHYKFCVDEWHRDERTITGDTGVNTLCITRD	105
QY	63	YNQINTLSSSTPGSRMMDVNDNFORTVTLSDGTSEGLRVSEAAIQISRCRVSEYL	122
DB	106	FDQINTLSSSTPGSRMMDVNDNFORTVTLSDGTSEGLRVSEAAIQISRCRVSEYL	165
QY	123	NLHFCYLLPDSGKVIALDNLVPKQSFHILHEQIGIPVAPLWDSFRGQVGLLSPDLFIL	182
DB	166	NGQGYDLLPDSGKVIALDNLVPKQSFHILHEQIGIPVAPLWDSFRGQVGLLSPDLFIL	225
QY	183	ILRELETHGSLNLTREQLTHETISAWKEAKCTNGRNDNSQWRPQOHLVHATPYESIRIAV	242
DB	226	ILRELETHGSLNLTREQLTHETISAWKEAKCTNGRNDNSQWRPQOHLVHATPYESIRIAV	285
QY	243	KLLQNGISTVPVIVSSSDSGSEFOLLHLASLSGILKICICRYFKNSTGNLPILNQPVCISIP	302
DB	286	KLLQNGISTVPVIMFSSSPDSYFOLLHLASLSGILKICICRYFKNSTGNLPILNQPVCISIP	345
QY	303	LGSWVPKIGDINSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNSLDTYSRSDITA	362
DB	346	LGTWVPKIGDINSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNSLDTYSRSDITA	405
QY	363	LAKDKVTHRLDEMTTHQALQLQDANTPPGFNGQRCQCLRSDDLKVMERLANP	420
DB	406	LAKDKVTHRLDEMTTHQALQLQDANTPPGFNGQRCQCLRSDDLKVMERLANP	463

RESULT 5

US-10-425-114-69679
; Sequence 69679, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)/B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69679
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLM017174G09_FLI.pep
US-10-425-114-69679

Query Match 76.9%; Score 1802; DB 12; Length 365;

Best Local Similarity 96.4%; Pred. No. 1.6e-170;

Matches 352; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY	87	NFQRTVTLSDGTSEGLRVSEAAIQISRCRVSEYLMLHTCYDILLPDSGKVIALDNLVP	146
DB	1	NFQRTVTLSDGTSEGLRVSEAAIQISRCRVSEYLMLHTCYDILLPDSGKVIALDNLVP	60
QY	147	KQSPHILHEQIGIPVAPLWDSFRGQVGLLSPDLFILILRELETHGSLNLTREQLTHETISA	206
DB	61	KQSPHILHEQIGIPVAPLWDSFRGQVGLLSPDLFILILRELETHGSLNLTREQLTHETISA	120
QY	207	WKEAKCTNGRNDNSQWRPQOHLVHATPYESIRIAVLLQNGISTVPVIVSSSDSGSPQ	266
DB	121	WKEAKCTNGRNDNSQWRPQOHLVHATPYESIRIAVLLQNGISTVPVIVSSSDSGSPQ	180
QY	267	LLHLASLSGILKICICRYFKNSTGNLPILNQPVCISIPGSLWPKIGDINSRPLAMLRPNAS	326
DB	181	LLHLASLSGILKICICRYFKNSTGNLPILNQPVCISIPGSLWPKIGDINSRPLAMLRPNAS	240
QY	327	LSSALNMLVQAGVSSIPVDDNSLDTYSRSDITAKDKVYTHRLDEMTTHQALQLG	386
DB	241	LSSALNMLVQAGVSSIPVDDNSLDTYSRSDITAKDKVYTHRLDEMTTHQALQLG	300
QY	387	QDANTPPGFNGQRCQCLRSDDLKVMERLANPQVRVFFIVEAGSKRVEGIIISLSDIPK	446
DB	301	QDANTPPGFNGQRCQCLRSDDLKVMERLANPQVRVFFIVEAGSKRVEGIIISLSDIPK	360
QY	447	FLLSL 451	
DB	361	FLLSL 365	

RESULT 6

US-10-424-599-187700
; Sequence 187700, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)/B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187700
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140506C.1.pep
US-10-424-599-187700

Query Match 65.2%; Score 1527.5; DB 12; Length 493;

Best Local Similarity 66.6%; Pred. No. 6.3e-143;

Matches 301; Conservative 58; Mismatches 82; Indels 11; Gaps 5;

QY 3 EHLPMSPTEGCTVFQALCSLSPGHEHYKFFVDGEMRDERQCTISGEFGIVNTL-YLTR 61
DB 46 ELLQMSPEVGCTVFQVHSHLLPGYHQYKFFVDGEMRDERQCTISGEFGIVNTL-YLTR 105
QY 62 EYNOINTLSSPSTPGSRMNDVDNENFQRTVTLSDGTVSEGTLRVSEAAIQISRCRVSEY 121
DB 106 DPNVIFVL--PPDVASGNSMDVDNDAFRXARLUDTGILSEVLPRLSDTDVQISQRISAF 163
QY 122 LNLHTCYDLLPDSGKVALDINLFPVKOSPHILHGGIPVAPLWDSFRGQFVGLLSPDLFI 181
DB 164 LSSHAVELLPESGKVALDINLFPVKOSPHILHGGIPVAPLWDSFRGQFVGLLSPDLFI 223
QY 182 LILBELTHGNSLTEROLETHITISAWKEAK-----RQNGRNDSONRQOHLVHATPYESL 237
DB 224 LILBELHNGNSLTEROLETHITISAWKEAK-----RQNGRNDSONRQOHLVHATPYESL 280
QY 238 RDIKVLKLONGISTVPIVYSSSDGSPFQLHLHSLGILKICRYFKNSGTGNLPILNQP 297
DB 281 KDIAMKILQKESVTPVPIHSSSEDSAPFQLHLHSLGILKICRYFKNSGTGNLPILNQP 340
QY 298 VCSIPGSGWPKIGMNSRPLAMLRPNASLSALMLVQAGVSSIPVDDNDLSDTYSR 357
DB 341 ICAIPVGWPKIGESNRPLAMLRPNASLSALMLVQAGVSSIPVDDNDLSDTYSR 400
QY 358 SDITLAKDKVYTHVRLEDMTHIQALQIGODANTPPGPFNGORCOMCLRSPLIKVMERL 417
DB 401 SDITLAKDKVYTHVRLEDMTHIQALQIGODANTPPGPFNGORCOMCLRSPLIKVMERL 459
QY 418 ANPGVRVFIIVEAGSKRVEGIIISLSDIPKFL 449
DB 460 ANPGVRVFIIVEAGSKRVEGIIISLSDIPKFL 491

RESULT 7
US-10-425-114-58697
; Sequence 58697, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jiongong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53333)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58697
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700575557_F.II.pap
US-10-425-114-58697

Query Match 63.7%; Score 1492; DB 12; Length 302;
Best Local Similarity 96.7%; Pred. No. 9.9e-140;
Matches 292; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 150 PHILHEQIPVAPLWDSFRGQFVGLLSPDLILRELETHGNSLTEROLETHITISAWKE 209
DB 1 PHILHEQIPVAPLWDSFRGQFVGLLSPDLILRELETHGNSLTEROLETHITISAWKE 60
QY 210 AKRQTGRNDSONRQOHLVHATPYESLURDIKVLKLONGISTVPIVYSSSDGSPFQLH 269
DB 61 AKRQTGRNDSONRQOHLVHATPYESLURDIKVLKLONGISTVPIVYSSSDGSPFQLH 120
QY 270 LASLSGILKICRYFKNSGTGNLPILNQPVCSIPGSGWPKIGMNSRPLAMLRPNASLS 329

DB 121 LASLSGILKICRYFKNSGTGNLPILNQPVCSIPGSGWPKIGMNSRPLAMLRPNASLS 180
QY 330 ALNMLVQAGVSSIPVDDNDLSDTYSRSDITLAKDKVYTHVRLEDMTHIQALQIGODA 389
DB 181 ALNMLVQAGVSSIPVDDNDLSDTYSRSDITLAKDKVYTHVRLEDMTHIQALQIGODA 240
QY 390 NTPGFFNGORCOMCLRSPLIKVMERLAMPGRVFIIVEAGSKRVEGIIISLSDIPKFL 449
DB 241 NTPGFFNGORCOMCLRSPLIKVMERLAMPGRVFIIVEAGSKRVEGIIISLSDIPKFL 300
QY 450 SL 451
DB 301 SL 302
RESULT 8
US-10-424-599-222531
; Sequence 222531, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 222531
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_42975C.1.pap
US-10-424-599-222531

Query Match 63.5%; Score 1487; DB 12; Length 461;
Best Local Similarity 63.6%; Pred. No. 6.2e-139;
Matches 287; Conservative 65; Mismatches 75; Indels 24; Gaps 4;
QY 4 HLPMSPIEGCTVFQALCSLSPGHEHYKFFVDGEMRDERQCTISGEFGIVNTLTYREY 63
DB 29 HLPMSPIEGCTVFQALCSLSPGHEHYKFFVDGEMRDERQCTISGEFGIVNTLTYREY 88
QY 64 NQINTLSSPSTPGSRMNDVDNENFQRTVTLSDGTVSEGTLRVSEAAIQISRCRVSEYN 123
DB 89 DILPSILNTETPG-RSHMEVDNMEANPRMTVSD-----LEVSRRHRSVFLS 133
QY 124 LHTCYDLLPDSGKVALDINLFPVKOSPHILHGGIPVAPLWDSFRGQFVGLLSPDLFI 183
DB 134 THTSYDLLPESGKVALDINLFPVKOSPHILHGGIPVAPLWDSFRGQFVGLLSPDLFI 193
QY 184 LRELETHGNSLTEROLETHITISAWKEAKRQNGRNDSONRQOHLVHATPYESL 238
DB 194 LKELGNHSSNLTEROLETHITISAWKEAKRQNGRNDSONRQOHLVHATPYESL 249
QY 239 DIAVKLQNGISTVPIVYSSSDGSPFQLHLHSLGILKICRYFKNSGTGNLPILNQPV 298
DB 250 DIAVKLQNGISTVPIVYSSSDGSPFQLHLHSLGILKICRYFKNSGTGNLPILNQPV 309
QY 299 CSIPGSGWPKIGMNSRPLAMLRPNASLSALMLVQAGVSSIPVDDNDLSDTYSR 358
DB 310 ASIPLGTWPKIGMNSRPLAMLRPNASLSALMLVQAGVSSIPVDDNDLSDTYSR 369
QY 359 DITLAKDKVYTHVRLEDMTHIQALQIGODANTPPGPFNGORCOMCLRSPLIKVMERL 418
DB 370 DITLAKDKVYTHVRLEDMTHIQALQIGODANTPPGPFNGORCOMCLRSPLIKVMERL 429
QY 419 NPGVRVFIIVEAGSKRVEGIIISLSDIPKFL 449
DB 430 NPGVRVFIIVEAGSKRVEGIIISLSDIPKFL 460


```
RESULT 9
US-10-425-114-54796
; Sequence 54796, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Lib, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54796
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GWFLMINSOY057G02_FLI.pep
US-10-425-114-54796
Query Match 56.6%; Score 1325; DB 12; Length 368;
Best Local Similarity 69.2%; Pred. No. 6.1e-123;
Matches 256; Conservative 52; Mismatches 54; Indels 8; Gaps 3;
QY 84 DNEPQRTVLTSGTSEGLTSEAAIQISRCRWSEYLNLTCTCYLLPDSGKVALDIN 143
DB 1 DNDAFERMAELTDTSEVLPRISDTDVQISQRISAFLSHTAYELLPESGKVALD 60
QY 144 LPVKQSFILHEQGIPVAPLWDSFRGQVGLSPDLFILRELETHGSLTBEQLETH 203
DB 61 LPVKQAFILHEQGVFMAPLWDFCKGFVGLSASDFILRELGNHGSNLTBELETH 120
QY 204 ISAWKEAK-----RQTGRNDSQWRPQOHLVHATPYESLRDIANKLQNGISTVPIYSS 259
DB 121 ISAWKEGKSYLNQNNHGHTAFSR---CFIHAGPIVNDKIDKAMKILQKEVSTVPIHSS 177
QY 260 SDGSPFOLLHLASGLTKICRYFKNSTGNLPINOPVCSIPGLSVPKIGDINSRPLA 319
DB 178 EDASFPOLLHLASGLTKICRYFHCSSSLPVQLPICAIPWGWPKIGESNRPLA 237
QY 320 MLRPNASLSALNMLVOAGVSSPIVDDNDSLLDTYSRSDITALAKDKVTHVRLDEMTI 379
DB 238 MLRPTASLASALNMLVOAQVSSPIVDDNDSLLDIYCRSDITALAKNRAYTHINLDEMT 297
QY 380 HQALQLOGDANTPEFFNGQRCOMCLRSDDLKVMERLANPGVRRVIVEAGSKRVEGII 439
DB 298 HQALQLOGDAYSPYE-LRSQRCOMCLRSDDLKVMERLANPGVRRVIVEAGSKRVEGIV 356
QY 440 SLSDIEKFL 449
DB 357 SLSDIEKFFI 366
RESULT 10
US-10-424-599-142943
; Sequence 142943, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 142943
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..{299}
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100095C.1.pep
US-10-424-599-142943
Query Match 39.3%; Score 920; DB 12; Length 299;
Best Local Similarity 62.1%; Pred. No. 1.1e-82;
Matches 187; Conservative 32; Mismatches 36; Indels 46; Gaps 5;
QY 182 LILRELETHGSLNTEEGLETHTISAWKEAKQTNGRNDSQWRP-QOHLVHATPYESLRDI 240
DB 10 LLKHKLGNHGSNLTBELETHTISAWKGGK-----WTGFTQCFIRAGPYDNKEI 59
```

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 142943
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..{263}
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100095C.1.pep
US-10-424-599-142943
Query Match 39.4%; Score 923.5; DB 12; Length 263;
Best Local Similarity 69.9%; Pred. No. 3.8e-83;
Matches 186; Conservative 31; Mismatches 34; Indels 15; Gaps 5;
QY 186 ELETHGSLNTEEGLETHTISAWKEAKQTNGRNDSQWRP-QOHLVHATPYESLRDI 244
DB 7 QLGNHGSNLTBELETHTISAWKGGK-----WTGFTQCFIRAGPYDNKEI 56
QY 245 LONGISTVPIYSSSDGSPOLLHLASLSGLKICRYFKNSTGNLPILNQPVCISPLG 304
DB 57 LQHGISTVPIIH--SEDSPPOLLHLASLSGLKICRYFKNSTGNLPILNQPVCISPLG 114
QY 305 SWPKIGDLNSRPLMLRPNASLSALNMLVOAGVSSPIVDDNDSLLD-TYSRSDITAL 363
DB 115 TWPKIGESNRPLMLRPNASLSALNMLVOAGVSSPIVDDNDSLLDIYCRSDITAL 174
QY 364 AKDKVYTHVRLDEMTIHQALQLOGDANTPEFFNGQRCOMCLRSDDLKVMERLANPGVR 423
DB 175 AKDRTYTHINLDEMTVHQALQLOGDSYNTYE-LSCQRCOMCLRTDSLHKVMERLANPGVR 233
QY 424 RVFIVEAGSKRVEGILSLDIEKFL 449
DB 234 RLVIIVEAGSKRVEGIIALSDFNFFL 259
RESULT 11
US-10-424-599-142947
; Sequence 142947, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 142947
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..{299}
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100095C.1.pep
US-10-424-599-142947
Query Match 39.3%; Score 920; DB 12; Length 299;
Best Local Similarity 62.1%; Pred. No. 1.1e-82;
Matches 187; Conservative 32; Mismatches 36; Indels 46; Gaps 5;
QY 182 LILRELETHGSLNTEEGLETHTISAWKEAKQTNGRNDSQWRP-QOHLVHATPYESLRDI 240
DB 10 LLKHKLGNHGSNLTBELETHTISAWKGGK-----WTGFTQCFIRAGPYDNKEI 59
```


RESULT 14
US-10-437-963-125316
; Sequence 125316, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 125316
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(166)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27972C.1.pep
US-10-437-963-125316

Query Match 33.2%; Score 778; DB 16; Length 166;
Best Local Similarity 92.1%; Pred. No. 5.9e-69;
Matches 152; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 286 NSTGNLPILNQPVCSTPLGSGWPKIGDLSNRPLAMLRPNASLSALNMLVQAGVSSPIW 345
Db 2 NSQGNLPILSQVCTPLGTWPKIGDPNRPPLAMLRPNLSALNMLVQAGVSSPIW 61

Qy 346 DENDSLDTSRSDITALAKDKVYTHVRLDEMTIHOALQIQDANTPFQFNGQRCQACL 405
Db 62 DENDSLDTSRSDITALAKDKVYTHVRLDEMTIHOALQIQDANTPFQFNGQRCQACL 121

Qy 406 RSPDLKWMERLANPGVRRVFIIVEAGSKRVEGIISLSDIFKLLS 450
Db 122 RSPDLKWMERLANPGVRRVFIIVEAGSKRVEGIISLSDIFKLLS 166

RESULT 15
US-10-425-114-50202
; Sequence 50202, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50202
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700799332_FJI.pep
US-10-425-114-50202

Query Match 32.5%; Score 761; DB 12; Length 477;
Best Local Similarity 38.8%; Pred. No. 1.6e-66;
Matches 173; Conservative 85; Mismatches 152; Indels 38; Gaps 8;

Qy 13 CP-----TVPQATCSLSPGHIHYKFFVDEGWRHDEROPTTSGEPGIVNTLYLTREYQIN 67
Db 56 CPMGLVGAFFVFDLPPGVYQVRFVLDGVYRCDETKPFVRDEYGLISNEVLENNVQPV 115

Qy 68 TISSPSTPGSRMMVDNENFQRTVTLSDGTV-----SEGLRVSEAAIQISRCRVSEYL 122
Db 116 VOPEPSIRGT--NMD-----KGTILKTMPEPSSQNPMSQCIATIRHVVSGIL 160

Qy 123 NLHTCYDILLPDSGKVIALDINLPKQSFHILHEQIPVAPLWDSFRGQFVGLLSPLJFIL 182
Db 161 LHTYIDVVLSSKLVLDLQVQAKFMHDEGLALVPLWDRQGTITGMLTASDFVL 220

Qy 183 ILRELETHGSLNTEEQLETHHTISAWKEAKQOTNGRNDQWPRPOOHLVHATPYESLRDIAV 242
Db 221 ILRKLQRNIQVIGNEE----PISAWKEAKLQFYGGPOGAAMQRPLIHVKDSDNLVDVAL 276

Qy 243 KLLQNGLSTVPVLYSSSSDGSFPOLLHLASLSGLIKCICRYFKNSTGNLPILNQPVCSIP 302
Db 277 TIIRNEISSVP-IFKCMADSSGVPEFLNLTATQGLKFLCSKIQBEAEGCSLHQLLSIP 335

Qy 303 LGSWVPKIGDLSNRPLAMLRPNASLSALNMLVQAGVSSPIVDDNDLSLDTYSRSDITA 362
Db 336 IGTWSPHTGRSSRQLRTLLSSPLNTCLDILLQDRYSSPIVDDNGSLRDVYSLSDIWA 395

Qy 363 LAKDKVYTHVRLDEMTIHOALQIQDANTPFQFNGQRCQACLRSQDLKWMERLANPG 421
Db 396 LAKNDVYARIELEQVTVQNALDVQYQV-----HGRROCHTCLQITLLEVLGLSIPG 448

Qy 422 VRRVFIIVEAGSKRVEGIISLSDIFKLL 449
Db 449 VRRLVWIEQSTREVEGIISLRDVFLL 476

Search completed: July 7, 2004, 17:53:56
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2004, 05:16:54 ; Search time 5618 Seconds
(without alignments)
3479.478 Million cell updates/sec

Title: US-09-857-525C-2
Perfect score: 2342
Sequence: 1 TREHLPMSPIEGPTVFQAI.....SKRVEGIISLSIDIFKFLSL 451

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US9857525/runat_07072004_161018_293/app_query.fasta_1.647
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=oct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9857525@cgn_1_1_3731@runat_07072004_161018_293 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_other.*
33: em_hgt_mus.*
34: em_hgt_pln.*
35: em_hgt_rnd.*
36: em_hgt_mam.*
37: em_hgt_vrt.*
38: em_sy.*
39: em_hgtc_hum.*
40: em_hgtc_mus.*
41: em_hgtc_other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2315	98.8	2156	8	AF276086 Zea mays
2	2218	94.7	2247	8	AF276085 Zea mays
3	2085	89.0	2145	8	AK102647 Oryza sat
4	2085	89.0	2565	8	AK111501 Oryza sat
5	2031	86.7	1632	8	BT009111 Triticum
6	1580	67.5	1464	8	BT000625 Arabidops
7	1580	67.5	2090	8	AF250335 Arabidops
8	1580	67.5	2301	8	AF439826 Arabidops
9	1481.5	63.3	1887	8	AY247268 Medicago
10	1171.5	50.0	92392	8	AC128647 Oryza sat
11	1171.5	50.0	145115	8	AC096690 Oryza sat
12	1068	45.6	982	8	BT009623 Triticum
13	1003.5	42.8	110514	8	AC00106 Sequence
14	798.5	34.1	1923	8	AK121122 Oryza sat
15	682	29.1	139089	2	AC146571 Medicago
16	661.5	28.2	107857	2	AC146552 Medicago
17	380.5	16.2	3067	5	BC060444 Xenopus l
18	379.5	16.2	2232	10	AY348865 Rattus no
19	377.5	16.1	2192	10	AY348864 Mus muscu
20	377.5	16.1	2194	9	AF087875 Homo sapi
21	374	16.0	3132	10	BC015283 Mus muscu
22	373.5	15.9	1167	9	AB025580 Homo sapi
23	373	15.9	987	9	BT007127 Homo sapi
24	373	15.9	987	12	BT008166 Synthetic
25	373	15.9	1435	6	AK139104 Sequence
26	373	15.9	2062	9	HA249976 Homo sapi
27	373	15.9	2203	9	BC020540 Homo sapi
28	373	15.9	2223	6	AX877755 Sequence
29	373	15.9	2223	6	BD156831 Primer fo
30	373	15.9	2223	9	AK001887 Homo sapi
31	371	15.8	3427	5	BC059181 Datio rer
32	367	15.7	996	9	BT007345 Homo sapi
33	367	15.7	1578	6	AR270550 Sequence
34	367	15.7	1578	9	HSU42412 Human 5' -AM
35	367	15.7	1668	9	BC000358 Homo sapi
36	366	15.6	3210	3	AY084138 Drosophil
37	363.5	15.5	2082	3	AF094764 Drosophil
38	363.5	15.5	3157	3	AY166758 Drosophil
39	363.5	15.5	3199	3	AY166753 Drosophil
40	363.5	15.5	3318	3	AY166757 Drosophil
41	363.5	15.5	3679	3	AY166756 Drosophil
42	363	15.5	2577	3	BT001510 Drosophil
43	363	15.5	2684	3	AY166755 Drosophil
44	363	15.5	2687	3	BT003249 Drosophil
45	363	15.5	2871	3	AY166754 Drosophil

ALIGNMENTS

RESULT 1

```

AF276086      2156 bp mRNA linear PLN 13-AUG-2002
LOCUS        Zea mays protein kinase AKINbetagama-2 mRNA, complete cds.
DEFINITION   AF276086
ACCESSION   AF276086
VERSION      AF276086.1 GI:11139547
KEYWORDS     Zea mays
SOURCE       Zea mays
ORGANISM     Zea mays
              Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 2156)
AUTHORS      Lumberreras, V., Alba, M.M., Kleinow, T., Koncz, C. and Pages, M.
TITLE        Domain fusion between SNPI-related kinase subunits during plant
              evolution
JOURNAL      EMBO Rep. 2 (1), 55-60 (2001)
MEDLINE      21146516
PUBMED       11252725
REFERENCE    2 (bases 1 to 2156)
AUTHORS      Lumberreras, V. and Pages, M.
TITLE        Direct Submission
JOURNAL      Submitted (07-JUN-2000) Genetica Molecular, CID, CSIC, Jordi Girona,
              18-26, Barcelona 08034, Spain
FEATURES     Location/Qualifiers
              1..2156
                 /organism="Zea mays"
                 /mol_type="mRNA"
                 /db_xref="taxon:4577"
                 /dev_stage="seedling"
                 /note="under water stress"
                 456..1946
                 /note="Snf-4 like"
                 /codon_start=1
                 /product="protein kinase AKINbetagama-2"
                 /protein_id="AA031752.1"
                 /db_xref="GI:11139548"
                 /translation="MESHGADSAHDAGVCVSSGGATVPRFVNPYGRKRVFVSGSF"
                 RSEHLPMPTEGCPVFQACISLPGHEIFKFFVGEWRHDERQPTISGEFIVNVL
                 YLTREYNQNTLSSPDPGSMNDVNFQVTLSDGTVEGTLRVSEAIIQISR
                 CRVSELYLTDCLDPSGKVIALDINLPVKQSFIHLHQGIPVAPLMDSPRGQVFG
                 LLSPLDFILIRLELYGSMLETSQLETHTSIAWKAQKQTNGRNDQMPQOHLVHA
                 TPVEDSRDIAVLQNGISVTPVYSSSDGSPOLLHLASLSGLKLCICRYFKNSTG
                 NFPIINQPVCSIFLGSWPKIGLNSRPFLAMLRPNASLSALANLWQAGVSSIPVDD
                 KNSLDTYGRSDITALAKAYTHYTRIDEMTHIQALQLQGDATPPGFNGQRQCMCL
                 RSDPLLKWERLANPGRVRVFIIVEAGSKRVEGIISUSIDIFKFLSL"
CDS           1..2156
              83 ValAspAsnGluAenPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
              837 GTGGATAATGAATAATTTCAACGACGTTACGTTGTGACATGTTACCGTTTCAGAGGT 896
              103 ThrLeuArgValSerGluAlaLalleGlnIleSerArgCysArgValSerGluTyrLeu 122
              897 ACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGCGCTGTTCTGAATATCTG 956
              123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
              957 AGTTTGCATACATGCTATGTTTACTCCCAAGATTCTGGCAAGGTTATTCCTCCAGACAT 1016
              143 AsnLeuProValLysGlnSerPheHisIleLeuHisGlnGlnGlyIleProValAlaPro 162
              1017 AATTTACCTGTGAAGCAATCTTCCATATTTCTCCATGAACAGGGGATTCCTGTAGCTCT 1076
              163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
              1077 CTCTGGGACTCATTCAGAGGTCATTTGTTGGTCTCTTACGCCCATTTGGATTTTCACTC 1136
              183 IleLeuArgGlnLeuGlnThrHisGlySerAsnLeuThrGluGlnGlnLeuGluThrHis 202
              1137 ATATTGGGAGCTAGAAACTCATGCTCGAATCTGACAGAGGAGGAGCTTGAACACAC 1196
              203 ThrIleSerAlaTyrLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTyr 222
              1197 ACTATATCTGCATGGAAAGAGGCTAAGCGCAACTAATCGAAGAAATGATGGTCAGTGG 1256
              223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGlnSerLeuArgAspIleAlaVal 242
              1257 CGACCGCAACAGCATCTAGTGTCATGCCGCCCTTATGAGTCTTTCAGGCGCATTCAGTA 1316
              243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly 262
              1317 AAGCTTTTGCAAAATGTCATTTCTACAGTCCAGTATTTATTTCTCATCATCATCATCAT 1376
              263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
              1377 TCATTCGGCAGTTATTCGATCTTCGATCATCTTCGAAATTTTGAATGATTTGTAGA 1436
              283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
              1437 TACTTCAAAAACCAACTGCTGTAATTTGCCCTTATTCGAACCAACCCAGTGTGCTCCATTCCG 1496
              303 LeuGlySerTyrValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
              1497 CTGGTTCTCTGGGTTCCGAAATTTGGTGTCTGAACAGTCTGTCATTTGGCTATGTTGCGA 1556
              323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
              1557 CCTAATGTCATCACCTTAGCTCTGCCCTTAAACATGTTGTTCAAGCTGAGTAAGCTCAATA 1616
              343 ProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
              1617 CCAATTCTGGATGCAACGACTCCCTCTCTTACACTTACTCTAGAGTGACATCACAGCC 1676
              363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
              1677 CTAGCAAAAGCCAGGCTTACACATGTTCCGCTGGATGGATGACCACTTCCACAGGCT 1736
              383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlnArgCysGln 402
              1737 TTGCAGCTTGGCAAGATGCCATACACCTTTTGGATTTCTTTAAACGCGCCAGAGATGCCAG 1796
              403 MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
              1797 ATGTGCTCTCGGCTGCTGATCTCTTCTGAGGTGATGAGGAGCTGGCTTAATCTCTGGGGTG 1856
              423 ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSer 442
              1857 CGGGGGGTGTTTCATTTGGGAAGCTGGGAGGCAACCTGTGGAGGCGCATCATCTAAGT 1916
              443 AspIlePheLysPheLeuLeuSerLeu 451

```

ORIGIN

```

Alignment Scores:
Pred. No.:      3 86e-189      Length:      2156
Score:          2315.00      Matches:      446
Percent Similarity: 99.55%      Conservative: 1
Best Local Similarity: 99.33%      Mismatches: 2
Query Match:     98.85%      Indels:      0
DB:              8              Gaps:        0

```

US-09-857-525C-2 (1-451) x AF276086 (1-2156)

```

QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
DB 597 GAACATTTGCGGATGTTCTCTATTGAAGCTGCCCACTGTATTTCAGGCTATTTCGAGT 656
QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu 42
DB 657 CTGTCTCTCGGATTCACGAGTACAAATTCCTTTGGAGCGGGAATGCGGCATGATGAG 716
QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
DB 717 CGTCAACTACCATATCTGGGAGTTTGGCATAGTTTAAACACTTTTACTTGACAAGGAA 776
QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
DB 777 TATAACCAATATAACACCTTATCAAGTCCAAGCACACCTGGAAGCAGGATGAACATGAT 836

```


Db 1900 CGCGCGGTGTTTCATAGTAGAAGCTGGGACAAACGTTGTGGAGGTATTATATCATCTAGT 1959
 QY 443 AspilephelysPheLeuLeuSerLeu 451
 Db 1960 GATATTTTCAAGTCTCTGTGAGCTTG 1986

AKI02647 2145 bp mRNA linear PLN 24-JUL-2003
 Oryza sativa (japonica cultivar-group) cDNA clone:J033100L02, full
 insert sequence.

AKI02647
 VERSION AKI02647.1 GI:32987856
 FLI CDNA: CAP trapper.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota: Viridiplantae; Streptophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
 1
 The Rice Full-length cDNA Consortium, National Institute of
 Agrobiological Sciences Rice Full-length cDNA Project Team,
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
 Ohtsuka, T., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group, Ohtsuka, T., Suzuki, Y., Tsunoda, Y.,
 Iida, Y., Sugano, S., Fujimura, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
 Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Miura, J.,
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., KIKEN,
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
 Kawai, J., Carninci, P., Hayatsu, N., Imotani, K., Arakawa, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayashi, K., Sakai, K., Sakazume, N., Sano, H.,
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oca, Y.,
 Saïto, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Yoshino, M., and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice
 Science 301: (5631), 376-379 (2003)

TITLE
 japonica rice

JOURNAL
 MEDLINE
 PUBMED
 22752273
 12869764

REFERENCE
 2 (bases 1 to 2145)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Maeda, H., Matsubara, C., Matsuyama, T., Miura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
 Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K.,
 Numasaka, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H.,
 Osato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
 Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yamada, H., Yamamoto, M.,
 Yasunishi, A., Yazaki, J., Yokomizo, S., and
 Yoshimura, A.
 Direct Submission
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
 Agrobiological Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
 Tel: 81-29-838-7007, Fax: 81-29-838-7007)
 This clone is one of the 28K full-length cDNA clones from japonica
 rice.

URL : <http://cdna01.dna.affrc.go.jp/cdna/>
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J.,
 Shikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, T., Shishiki, T., and
 Yamamoto, M.
 FAIS Genome Sequencing & Analysis Group: Ohtsuka, Y., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
 Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
 Yoshimura, A., Matsubara, C., and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaka, R., Ohno, M., Osato, N.,
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K.,
 Yasunishi, A., and Hayashizaki, Y.
 Location/Qualifiers
 1. 2145
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J033100L02"

FEATURES
 source
 Alignment Scores:
 Pred. No.: 2,24e-169 Length: 2145
 Score: 2085.00 Matches: 391
 Percent Similarity: 94.87% Conservative: 34
 Best Local Similarity: 97.28% Mismatches: 23
 Query Match: 89.03% Indels: 0
 DB: 8 Gaps: 0

US-09-857-525C-2 (1-451) x AKI02647 (1-2145)

QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 Db 204 GAACATTTGCCAATGCTCTCTCTTGAAGGTTGTCCACCGTTTTCAGGCCATATGCGAGT 263
 QY 23 LeuSerProGlyIleHisGluTyrIlysPhePheValAspGlyGluTrpArgHisAspGlu 42
 Db 264 TTGTCCTCCCTGGGATTCATCAGTACAAAGTTTGTGTGGATGGGAATGCGGCACCATGAG 323
 QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
 Db 324 CGGCAACCTTACCATAACAGGAGCATATGCTGTGGTTAACTTTATGCTTGTGCTAGTAGGAC 383
 QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 Db 384 TTTGACCAAGTAAATACAAATATTGAGCCCTTAGTACACCTTGAAGTAGGATGAACATGAT 443
 QY 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 Db 444 GTGGACACAGCAATTTTCAACGTAAGTCTCTCTTGTCTGTGTCGATGATTTTTCAGGAAGT 503
 QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
 Db 504 CCTCAGAGAAATTCAGAGGCGAGCTATACAGATCTCTAGGTGTGTGTAGCGGATTTCTG 563
 QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyIleValIleAlaLeuAspIle 142
 Db 564 AATGGCAAACTGGGTATGATTACTCCAGATTTCTGGCAAGTCTTGTCTCTAGACGTT 623

Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurinara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Osato, N.,
Oca, Y., Saichoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, X.,
Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, I., Tomaru, A., Ioya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

1. 2565
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:3994"
/clone="J01300N04"

ORIGIN

Alignment Scores:

Pred. No.: 2,83e-169 Length: 2565
Score: 2085.00 Matches: 392
Percent Similarity: 94.87% Conservative: 34
Best Local Similarity: 87.28% Mismatches: 23
Query Match: 89.03% Indels: 0
DB: 8 Gaps: 0

US-09-857-525C-2 (1-451) x AK111501 (1-2565)

Qy 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
Db 614 GAACATTGGCCAAATGTCCTGTTGAAAGGTTGTCCACCGCTTTCAGGCCCATATGCGAGT 673
Qy 23 LeuSerProGlyIleHisGluTyrlsPhePheValAspGlyGluTroArgHisAspGlu 42
Db 674 TTGTGCCCTGGATTCTACATGACGATCAAGTTTGTGTGATGGGGAATGGCGCACGATGAG 733
Qy 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrlsLeuThrArgGlu 62
Db 734 CGGCAACTACCAATACAGGAGCATGCTGCTGCTTAACACTTATGCTTGACTAGGGAC 793
Qy 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
Db 794 TTTGACCAGATAAATACAAATATTGAGCCCTAGTACACCTGGAGTAGGATGAACATGGAT 853
Qy 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
Db 854 GTGCAACAGCAGCAATTTTCAACGACGATGTTCTTGTCTGATGGCATTATTCAGAAAGGT 913
Qy 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrlsLeu 122
Db 914 CTTCAGAGAATTCAGAGGCGACTATACAGATCTAGTCTGCTGCTAGCGGATTTCTG 973
Qy 123 AsnLeuHisThrCysTyrlsAspLeuLeuProAspSerGlyIleValIleAlaLeuAspIle 142
Db 974 AATGGCAAACTGGGTATGATTACTCCCAAGATCTGGCAAGGTCATTTGCTCTAGACGTT 1033
Qy 143 AsnLeuProValIleGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
Db 1034 AATTGGCTGTGAAGCAATCTTTTCATATTTCTCATGACAGGGAATTCCTGTGGCAGCT 1093
Qy 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db 1094 CTATGGGATTCATTGAGGGGCGAGTTTGTGGCCCTTTGAGCCCACTGGATTTTATATCTC 1153
Qy 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHis 202
Db 1154 ATATTGAGAGCTGGAACTCATGGCTCCAAATCTGACAGAAAGACGCTTGAACACAT 1213

Qy 203 ThrIleSerAlaTrpLysGluAlaIleYsArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
Db 1214 ACTATATCTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1273
Qy 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrlsGluSerLeuArgAspIleAlaVal 242
Db 1274 AGGCAATATCATTATGATTCATGCCACCTTATGAAATCCCTGAGGGAATTCGTATG 1333
Qy 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrlsSerSerSerAspGly 262
Db 1334 AAGATATCTGCAAAACGGGTGTTCTACGCTTCCAAATATGTTTCTCTCATCAGGATGSC 1393
Qy 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLeuCysIleCysArg 282
Db 1394 TCATATCCAAATTTGTTGATCTTGTCTCCCTTCTGGAATTTTGAATGATTTGTAGA 1453
Qy 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
Db 1454 TATTTTAAATTTCTCAAGGTAATTTACCTAATTTTGGCCCACTGATGACAAATTCCT 1513
Qy 303 LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
Db 1514 CTGGGTACCTGGGTTCCAAATTTGGTGATCTTAATGGCCGCTCCATTTGGTATGTTGGG 1573
Qy 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerIle 342
Db 1574 CTTAACACATCTCTTAGCGCTGCCCTGAATTTGCTGGTTCAGCTGGTGTGAGCTCAATA 1633
Qy 343 ProIleValAspAspAsnAspSerLeuLeuAspThrTyrlsArgSerAspIleThrAla 362
Db 1634 CCAATTTGGATGACAAATGACCTCACTCTTGACACATATTCAGAAAGTGACATCAGCT 1693
Qy 363 LeuAlaIleAspLysValTyrlsHisValArgLeuAspGluMetThrIleHisGlnAla 382
Db 1694 CTTGCAAAAGCAAGGTGTACACACATTCGCTTAGATGAGACCATTCATCAGGCT 1753
Qy 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlyArgCysGln 402
Db 1754 CTGAGCTTGACAGGATGCAATTTGCCCTTTGGATTTCTTTACGGGCAAGATGTGAG 1813
Qy 403 MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
Db 1814 ATGTGCTCCGCTCTGACACTTTGTGAAAGTGATGAGGAGGATAGCTAACTCTGGG 1873
Qy 423 ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSer 442
Db 1874 CGCGCTGTTTTCATTTGTGAGGCTGTGAGCAACGCTGTGGAGGGGCAATATCATCTGAGT 1933
Qy 443 AspIlePheLysPheLeuLeuSer 450
Db 1934 GATATTTTCAAGTCTTGTGAGC 1957

RESULT 5

BT009111

LOCUS

DEFINITION

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

USA

BT009111 1632 bp mRNA linear PLN 20-JUN-2003
Triticum aestivum clone w11.pk0008.b3:fls, full insert mRNA

BT009111 GI:32128662

F1 CDNA.

Triticum aestivum (bread wheat)

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; triticum.

1 (bases 1 to 1632)

Tingey, S.V., Wolters, P., Powell, W., Dolan, M., Miao, G.-H.,

Caraher, N.R., Hanafey, M.K. and Hainey, C.F.

Direct Submission

Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and

Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,

USA

FEATURES

Location/Qualifiers
 1. .1632
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="w11.pk0008.b3.fis"

ORIGIN

Alignment Scores:

Pred. No.: 6,84e-165 Length: 1632
 Score: 2031.00 Matches: 386
 Percent Similarity: 93.74% Conservative: 33
 Best Local Similarity: 86.35% Mismatches: 26
 Query Match: 86.72% Indels: 2
 DB: 8 Gaps: 1

US-09-857-525C-2 (1-451) x BT009111 (1-1632)

QY 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerLeu 23
 DB 2 CACGAGCCGATGCTCTCGAGTTCGAGGTTGCCCACTGTATTTCAGCTATCTGCACTTA 61
 QY 24 SerProGlyIleHisGluTyrPhePheValAspGlyGluCpArgHisAspGluArg 43
 DB 62 CCTCCAGGAGTTATCAGTACAAAGTTCAATGTGCACGCGGCGAGTGGAGGCATGACGAGGG 121
 QY 44 GlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGluTyr 63
 DB 122 CAACCTACTATAACTCGAGAGTATGGGTGGTAAACACTTATATCTGACAGGGAATTT 181
 QY 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83
 DB 182 GACCACATTAATCTGTACCGAGCCCACTACCTCGGAGCAGG-----ATGGATGTG 235
 QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
 DB 236 GACAGTGACAGTTTCAACGAATGGGTTCGTTCGGATGGTGCCTTCAGGAAGGTCTCT 295
 QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsn 123
 DB 296 CCAAGAATCTCAGAGGCTGTATACAGATCTCTAGGTGTCTGTGTCTGAGTATCTGAAAT 355
 QY 124 LeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
 DB 356 GCGCATACAGGCTATGACCTACTACAGATCTCTGGAAGGTCATTGCTCGGCATTAAT 415
 QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyLeuProValAlaProLeu 163
 DB 416 TTACCTGTGAAGCAATCTTCCATATTCCTCATGACAGAGGAGTCTCTGTGGCTCTCTGT 475
 QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183
 DB 476 TGGGATTCATTCAGGGGTCAAGTTGTGGCTTCCTGAGCCACCTGGATTTTATCTATCA 535
 QY 184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHisThr 203
 DB 536 TTCAGAGAGCTGGAACCTCATGGCTCAAACTGACAGAGGAACAGCTTGAACACACACT 595
 QY 204 IleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArg 223
 DB 596 ATATCTCGGTGAAGAGGCTAAGCGGCAAACTTATGGAAGAATGATGGACAACTTAGA 655
 QY 224 ProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLys 243
 DB 656 TCAATCAGCATCTAGTGTGATGCCACCCCTTATGAAATCCCTGAGGGGTATTCGCATGAAA 715
 QY 244 LeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAspGlySer 263
 DB 716 ATACTCGAAACTGCAATTTCTACAGTCCCAATCATCTATTTCATCATCAGATGGATCG 775
 QY 264 PheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyr 283
 DB 776 TTTCGGCAGCTGTGTGCATCTTGCATCCGCTTTCAAGGAATTTTGAATGATCTGTAGATAC 835

QY 284 PheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeu 303
 DB 836 TTCAGAACTCCACTGGTAGTTTGGCGATTCTAAACCAACCACTATGCTCAATTCGCTG 895
 QY 304 GlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgPro 323
 DB 896 GGTACCTGGGTTCCAAAAATTCGTGAACCAATAGGTTCATCCATTCGCTATGTTGCGGCT 955
 QY 324 AsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIlePro 343
 DB 956 AATACATCTCTTAGCTTCGCCCTTAACCTTGTTCAGCTGAGTAGTTCATACCC 1015
 QY 344 IleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeu 363
 DB 1016 ATTGTGGATGATTAACGACTCGCTGATCGACACATCTCCAGAGTGCATCACAGCTCTA 1075
 QY 364 AlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeu 383
 DB 1076 GCGAAAGACAAAGCTTACACCCATATCCGCTAGATGAGATGACCATTCATCAGGCCCTG 1135
 QY 384 GlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheSerGlyGlnArgCysGlnMet 403
 DB 1136 CAGCTCGGCGACAGCGGAATTCACCTTTTGGACTTTTCAATGGTCAAGAGTCCAGATG 1195
 QY 404 CysLeuArgSerAspProLeuLysValMetSluArgLeuAlaAsnProGlyValArg 423
 DB 1196 TGCTCCAGTCTGACCTTTTCTGAGGTTATGAGAGATTGGCTAATCTCTGGGGTGGT 1255
 QY 424 ArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAsp 443
 DB 1256 CGCGTGTTCATCTCGGAGGCTGCGCAGCAGCGAGTGGAGGCGTAAATATCTCGTGGCGAC 1315
 QY 444 IlePheLysPheLeuLeuSer 450
 DB 1316 ATATTCAAGTTGCTGCTGAGC 1336

RESULT 6
 BT000625 1464 bp mRNA linear PLN 25-SEP-2002
 LOCUS Arabidopsis thaliana putative activator subunit of SNF1-related
 DEFINITION protein kinase SNF4 (At1g09020/F7G19_11; mRNA, complete cds.
 ACCSSION BT000625
 VERSION BT000625.1 GI:23308442
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 REFERENCE 1 (bases 1 to 1464)
 AUTHORS Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
 Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
 Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
 Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C.,
 Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W.,
 Theologis A. and Ecker J.R.
 TITLE Arabidopsis ORF clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1464)
 AUTHORS Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
 Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
 Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
 Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C.,
 Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W.,
 Theologis A. and Ecker J.R.
 TITLE Direct Submission
 JOURNAL Submitted (25-SEP-2002) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA': Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, P3EC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Banth, J., Bowser, L., Chan, M.M., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers
1. 1464
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/ecotyped="U14970"
/ecotype="Columbia"
/notes="This clone is in pUNI 51"
1. 1464
/notes="putative activator subunit of SNF1-related protein kinase SNF4"
/codon_start=1
/product="Atg39020/P7G19_11"
/protein_id="AA18191.1"
/db_xref="GI:23308443"
/translation="MFQSTLDSRNGSAAGQLPTFRFVMPYGGRRVFLSGSPTRWT EYVPMSELCPTFVQICNLPGYHOYKFPVGDGEMHDEHPVSGNGGVNTIFIT GDMVPAGFSPFELGRNMDVDFLTADPSQAVRMSGVDLERHRLSVLLSFR TAYDLPESEKVALDNLVVKQAFHLYEQGIPLELMDFGQFVGVLPLOFLI LRELGTHSLTEELTEETIAWKEGKAHISROYDSGRYPRLPVGVGVNDLKV ALKILQNAAPVLYESLDQSYFQPLHLSLGLKICRYFRHSSSLPILQDPI CSPIPLGVPRIGVYSGKPLATPHASLGSLALVQAEVSSI PVVDNDNDLSDIYS RSDITALKAKQVAGIHLDDMTVHQALQGDQSPVYGFIFNGQRCHMCLRSLSLVKM ERLANPGVRLRIVIVEAGSKRVEGILSLSDVPEQLLG."

CDS

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA': Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.)
The Salk, Stanford, P3EC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Banth, J., Bowser, L., Chan, M.M., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.)
Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.
Location/Qualifiers
1. 1464
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/ecotyped="U14970"
/ecotype="Columbia"
/notes="This clone is in pUNI 51"
1. 1464
/notes="putative activator subunit of SNF1-related protein kinase SNF4"
/codon_start=1
/product="Atg39020/P7G19_11"
/protein_id="AA18191.1"
/db_xref="GI:23308443"
/translation="MFQSTLDSRNGSAAGQLPTFRFVMPYGGRRVFLSGSPTRWT EYVPMSELCPTFVQICNLPGYHOYKFPVGDGEMHDEHPVSGNGGVNTIFIT GDMVPAGFSPFELGRNMDVDFLTADPSQAVRMSGVDLERHRLSVLLSFR TAYDLPESEKVALDNLVVKQAFHLYEQGIPLELMDFGQFVGVLPLOFLI LRELGTHSLTEELTEETIAWKEGKAHISROYDSGRYPRLPVGVGVNDLKV ALKILQNAAPVLYESLDQSYFQPLHLSLGLKICRYFRHSSSLPILQDPI CSPIPLGVPRIGVYSGKPLATPHASLGSLALVQAEVSSI PVVDNDNDLSDIYS RSDITALKAKQVAGIHLDDMTVHQALQGDQSPVYGFIFNGQRCHMCLRSLSLVKM ERLANPGVRLRIVIVEAGSKRVEGILSLSDVPEQLLG."

ORIGIN

Alignment Scores:
Pred. No.: 3,41e-126 Length: 1464
Score: 1580.00 Matches: 304
Percent Similarity: 81.78% Conservative: 64
Best Local Similarity: 67.56% Mismatches: 74
Query Match: 67.46% Indels: 8
DB: 8 Gaps: 5
US-09-857-525C-2 (1-451) x BT000625 (1-1464)
QY 3 GluHisLeuProMetSerProIleGlyCysProThrValPheGlnAlaIleCysSer 22
DB 133 GAGCATGTGCGCAATGTACCGCTCGAGCGCTGCCCTACTGTGTTTCAAGTCATTGCAAC 192
QY 23 LeuSerProGlyIleHisGluTyrllysPheValAspGlyGluTrpArgHisAspGlu 42
DB 193 TTGAGCCAGGATATCATCAGTATAAGTTTGTGTGATGGGAATGCGGCACAGTACG 252
QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrlLeuArgGlu 62
DB 253 CACCAACCATTTGTAAGCGGAATGGTGGAGTAGTGAATACATATTTATACTGGGCCA 312
QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
DB 313 ---GATATGTTCTGCTGCTGTTTATAGCCAGACACTGGGC---CGGTGCAATATGGAT 366

QY 83 ValAspAsnGluAsnPheGlnAcqThrValThrLeuSerAspGlyThrValSerGluGly 102
DB 367 GTGGAT---GATGTCCTTCGCGAAGCGCTGACCCGTC---CAGGAGCT 411
QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrlLeu 122
DB 412 GTTCTAGGATGTCGGGGTGAATTGCGAGCTTCTCGTCACCGCATATCGTTTATTG 471
QY 123 AsnLeuHisThrCysTyrlAspLeuProAspSerGlyIleValIleAlaLeuAspIle 142
DB 472 TCACCCCGCATGTCATATGAGCTCTCCCAAGTCGGGCAAGGTATTGCAATGGATGTT 531
QY 143 AsnLeuProValIleGlnSerPheHisIleLeuHisGlnGlnGlyIleProValAlaPro 162
DB 532 AATTTGCGATGAGCAGCAGCATTCATATATCTATGAGCAGGGAATCCCTTTGGCTCT 591
QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
DB 592 CTTTGGGACTTTCGAAAGGCAATTTGTTGGAGTCTCTGTCCTCCACTAGACTTCATTCTA 651
QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnGluLeuGluThrHis 202
DB 652 ATACTGAGAGAGCTTGGAACTCATGGATCAACATTGACAGGAGAGCTTGGAGCGCAC 711
QY 203 ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
DB 712 ACAATAGCAGCTCGAAAGAGGGAGGCTCATATTAGCCGCAATATGATGGAAGTGGG 771
QY 223 ArgPro---GlnGlnHisLeuValHisAlaThrProTyrlGluSerLeuArgAspIleAla 241
DB 772 AGACCATATCTAGGCCACTTGTTCAGCTTGGAGCCCTATGATAATCTGAAAGAGCTGCC 831
QY 242 ValIleLeuLeuGlnAsnGlyIleSerThrValProValIleTyrlSerSerSerSerAsp 261
DB 832 CTGAAATTTTGGCAAAACAGGTGGCTGCTTCGCTATCAGGCATATTAATGATGATGC 891
QY 262 GlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCys 281
DB 892 GGTTCATATCCGAGTTACTTCATCTGCTTCGCTATCAGGCATATTAATGATGATGC 951
QY 282 ArgTyrlPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
DB 952 AGATACCTTTAGACATCTGCTAGCTCTTACCTATCTCTTCAGCAACCCATTGTTCAAT 1011
QY 302 ProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu 321
DB 1012 CCCTCGGTAGTGGTCCCTAGATCGAGATCAAGTAGCAACCTCTCGTACATTG 1071
QY 322 ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341
DB 1072 AGACTCAGGCTCTTTGGGTTCGCGCTCGCATATTATTAGTTCAAGCTGGAAGTCAGTTCA 1131
QY 342 IleProIleValAspAsnAspSerLeuLeuAspThrTyrlSerArgSerAspIleThr 361
DB 1132 ATTCCGTAGTGGATGACACGACTCGCTTATTGACATATCTCTCGAAGTGATTAATCT 1191
QY 362 AlaLeuAlaLysAspLysValTyrlThrHisValArgLeuAspGluMetThrIleHisGln 381
DB 1192 GCTCTGGCTTAAGACAGGCATACGCAGATTCTCTTGATGACATGACAGTTCAACCAG 1251
QY 382 AlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys 401
DB 1252 CGCTACAGTTGGGCAAGATCGGACCGCGCTTATGGAATCTTCAACGGGCGAGAGATGT 1311
QY 402 GlnMetCysLeuArgSerAspProLeuLysValMetGluArgLeuAlaAsnProGly 421
DB 1312 CACATGTCTTGGCTCAGACTCTCTGTGAAAGTGTGGAGCGGTTCGGCAATCCAGGT 1371
QY 422 ValArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeu 441
DB 1372 GTAAGAGGCTGTGTATAGTGAAGCAGGAGCAACCTGTTGAAGGTATCATTTCACTG 1431
QY 442 SerAspIlePheLysPheLeuLeuSerLeu 451

```

Db      1432  ACAGATGTTTCAATTCCTGCTCGTCTT 1461
|||||:|||||:|||||:|||||:|||||
RESULT 7
AF250335      2090 bp      mRNA      linear      PLN 06-SEP-2000
LOCUS      Arabidopsis thaliana putative activator subunit of SNF1-related
DEFINITION      protein kinase SNF4 mRNA, complete cds.
ACCESSION      AF250335
VERSION      AF250335.1
KEYWORDS      GI:9965728
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2090)
Kleinow,T., Bhalerao,R., Breuer,F., Umeda,M., Salchert,K. and
KoncZ,C.
Functional identification of an Arabidopsis snf4 ortholog by
screening for heterologous multicopy suppressors of snf4 deficiency
in yeast
Plant J. 23 (1), 115-122 (2003)
20387008
10929106
2 (bases 1 to 2090)
Kleinow,T., Bhalerao,R. and KoncZ,C.
Direct Submission
AUTHORS      Kleinow,T., Bhalerao,R. and KoncZ,C.
TITLE      Submitted (29-MAR-2000) Genetic Principles of Plant Breeding,
Max-Planck-Institute for Plant Breeding, Carl-von-Linne-Weg 10,
KoeIn 50829, Germany
FEATURES
source
1..2090
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/map="1-8,7"
932..2080
/note="contains 4 cystathione-beta-synthase (CBS) domains;
heterologous multicopy suppressor of snf4 deficiency in
yeast"
/codon_start=1
/product="putative activator subunit of SNF1-related
protein kinase SNF4"
/protein_id="AAG10141.1"
/db_xref="GI:9965729"
/translation="MVPAGFSPETLGRSNMDDVFLRTADPSQERVPRMSGVDEL
SRHSVLLSTRTRAYELPESGKVIADVNLVPRKAFHLYEQIPLAPLWDFKGQFY
GVGLPDLFILIKELGTHGNSLTRELETHIAAMKEKAHISQYDGSGRPYRPLV
GVGPDNLKQVAKILQNKVAAVPVYVSLQDGSYPQLHLASLGLKICICYPFRHS
SSSLPILQPICSIPLGTWVPRGESSKPLATLRPHASIGSALALLVQAEVSSIPV
DNDSLIDYISRDITALAKDAXAQIHLDMTHQALQLQDASPYGIFNGQRCHM
CLRDSLIVKMERLANPVRRLVIVRGSKEVEGIIISLSDVFLQL"
CDS
ORIGIN
Alignment Scores:
Pred. No.:      5,43e-126      Length:      2090
Score:      1580.00      Matches:      304
Percent Similarity:      81.78%      Conservative:      64
Best Local Similarity:      67.56%      Mismatches:      74
Query Match:      67.46%      Indels:      8
DB:      8      Gaps:      5
US-09-857-525C-2 (1-451) x AF250335 (1-2090)
QY      3      GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
|||||:|||||:|||||:|||||:|||||
Db      749      GACGATGTCGAATGTCACCGCTCGAGGGTGCCTACTGTTTTCAGTCAATTCGCAAC 808
|||||:|||||:|||||:|||||:|||||
QY      23      LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluThrArgHisAsnGlu 42
|||||:|||||:|||||:|||||:|||||
Db      809      TTGACGCCAGGATATCATCAGTATAGTATTITGTTGATGGGAATGGCGGCAGATGAG 868
|||||:|||||:|||||:|||||:|||||
43      ArgGlnProThrIleSerGlyGluPheGlyValAsnThrLeuTyrLeuThrArgGlu 62
|||||:|||||:|||||:|||||:|||||
869      CACCAACCATTTGTAAGCGGAATGGTAGTAGTGAATACAATATTTATAACTGGGCCA 928
|||||:|||||:|||||:|||||:|||||
63      TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
|||||:|||||:|||||:|||||:|||||
929      ---GATATGGTTCTGCTGCTGTTTAGCCACAGACACTGGGC---CGTCCAAATATCGAT 982
|||||:|||||:|||||:|||||:|||||
83      ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
|||||:|||||:|||||:|||||:|||||
983      GTGGAT---GATGTCCTTCGCGACGCGCTGACCGCTCC-----CAGGAGACT 1027
|||||:|||||:|||||:|||||:|||||
103      ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
|||||:|||||:|||||:|||||:|||||
1028      GTTCTTAGGATGTCGCGGGTGTGATTGGAGCTTCTCGTCACGCATATCGGTTTATTG 1087
|||||:|||||:|||||:|||||:|||||
123      AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
|||||:|||||:|||||:|||||:|||||
1088      TCAACCCCGCACTGCATATAGCTGCTCCCAAGATCGGCAAGGTATTGCAATTGGATGTT 1147
|||||:|||||:|||||:|||||:|||||
143      AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIlePheValAlaPro 162
|||||:|||||:|||||:|||||:|||||
1148      AATTGCCAGTGAAGCAAGACATTCATTACTCTATGACGAGGAATCCCTTGGCTCT 1207
|||||:|||||:|||||:|||||:|||||
163      LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
|||||:|||||:|||||:|||||:|||||
1208      CTTTGGGACTTTGMAAAGGCCAATTTGTTGGAGTTCTTGCTCCACTAGACTTCATTCTA 1267
|||||:|||||:|||||:|||||:|||||
183      IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHis 202
|||||:|||||:|||||:|||||:|||||
1268      ATACTGAGAGAGCTTGGAACTCATGGATCAAACTTGACAGGAAGAGCTTGACAGCGAC 1327
|||||:|||||:|||||:|||||:|||||
203      ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
|||||:|||||:|||||:|||||:|||||
1328      ACAATAGCAGCTCGAAGAGGGGAGGCTCATATTAGCGGCATATGATGGAGTGGG 1387
|||||:|||||:|||||:|||||:|||||
223      ArgPro---GlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAla 241
|||||:|||||:|||||:|||||:|||||
1388      AGACCATATCTAGGCCACTTGTTCAGGTTGGACCTCATGATATCTGAAAGACGTTGCC 1447
|||||:|||||:|||||:|||||:|||||
242      ValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAsp 261
|||||:|||||:|||||:|||||:|||||
1448      CTGAAATTTTGCAAAACAAAGTGGCAGCTGTTCCGGTTATATATCTCTTTCCGAGAT 1507
|||||:|||||:|||||:|||||:|||||
262      GlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCys 281
|||||:|||||:|||||:|||||:|||||
1508      GGTTCATATCCGCGATTCTGCTGCTGCTATCAGGCATATTAAATGTTATATGC 1567
|||||:|||||:|||||:|||||:|||||
282      ArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
|||||:|||||:|||||:|||||:|||||
1568      AGATACCTTTAGACATTCGCTAGCTCTTTACCTATCTCTCAGCAACCCCAATTTGTTCA 1627
|||||:|||||:|||||:|||||:|||||
302      ProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu 321
|||||:|||||:|||||:|||||:|||||
1628      CCCTCGGTACGTGGTGGTCCCTAGATCGGAATCAAGTAGCAAACTCTTGCTACATG 1687
|||||:|||||:|||||:|||||:|||||
322      ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341
|||||:|||||:|||||:|||||:|||||
1688      AGACCTCAAGCTCTTTGGGTTCTGCGCTGCATTATAGTCAAGCTGAAGTCAGTCA 1747
|||||:|||||:|||||:|||||:|||||
342      IleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThr 361
|||||:|||||:|||||:|||||:|||||
1748      ATTCCTCGTAGTGGATGACAACGACTCGCTTATTGACATATACTCTCGAAGTGATATACT 1807
|||||:|||||:|||||:|||||:|||||
362      AlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGln 381
|||||:|||||:|||||:|||||:|||||
1808      GCTCTGCTTAAGACACAGGCATACGCACGATTCATCTTGTGACATGACATGTCACAG 1867
|||||:|||||:|||||:|||||:|||||
382      AlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys 401
|||||:|||||:|||||:|||||:|||||
1868      GCGCTACAGTTGGGGCAAGATGATGAGCCGCCCTTATGGAATCTTCAACGGCGCAGATGT 1927
|||||:|||||:|||||:|||||:|||||

```


QY	402	GlnMetCysLeuArgSerAppProLeuLeuLysVa.MetGluArgLeuAlaAsnProGly	421
	1928	CACATGTCCTTGGCTCAGACTCTCTGTGAAAGTGATGAGCGGTTGGCGAATCCAGGT	1987
Db			
QY	422	ValArgValPheLeuValGluAlaGlySerLysArgValGluGlyLeuLeuSerLeu	441
Db	1988	GTAAGGAGGCTGGTATAGTGGAGGAGGAGCAACGCTGTGAGGTATCAATTCAGT	2047
QY	442	SerAspLeuPheLysPheLeuLeuSerLeu	451
Db	2048	AGCATGTTTCCAAATTCCTGCTCGGCTT	2077
RESULT 8	AF439826	2301 bp	linear
LOCUS	Arabidopsis thaliana At1g09020/F7G19_11	mRNA	PLN 04-NOV-2001
DEFINITION	Arabidopsis thaliana		complete cds.
ACCESSION	A2439826	GI:16612254	
VERSION	F.1	CDNA.	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE	1	(bases 1 to 2301)	
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bowser, I., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.		
TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCE	2	(bases 1 to 2301)	
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bowser, I., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-OCT-2001) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
	The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J., Bowser, I., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.		
FEATURES	source		
	1..2301		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="1"		
	location/Qualifiers		
	1..2301		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="1"		
	misc_difference	1267	
	/note="compared to genome"		
	/replace="g"		
	2068..2301		
	3'UTR		
	ORIGIN		
	Alignment Scores:		
	Pred. No.:	6,17e-126	Length: 2301
	Score:	1580.00	Matches: 304
	Percent Similarity:	81.78%	Conservative: 64
	Best Local Similarity:	67.56%	Mismatches: 74
	Query Match:	67.46%	Indels: 8
	DB:	8	Gaps: 5
	US-09-857-525C-2 (1-451) x AF439826 (1-2301)		
QY	3	GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer	22
Db	736	GAGCATGTGCCAATGTCACCGCTCGAGGCTGCCTACTGTTTTCAGTCAATTCGCAAC	795
QY	23	LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluThrArgHisAspGlu	42
Db	796	TTGACGCCAGGATATCATCAGTATAAGTTTTTTGTGATGGGGAATGCGGCACCATGAG	855
QY	43	ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu	62
Db	856	CACCAACCATTTGAAGCGGAATGCTGGAGTAGTAGTGAATACATATTATTAATCGGCCA	915
QY	63	TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp	82
Db	916	---GATATGTTCTCTGCTGCTGTTTATGCCCAGAGACACTGGGC---CGGTGGAATATGGAT	969
QY	83	ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly	102
Db	970	GTGGAT---GATGTTCTCTCGGAACGCTGACCCGCTC-----CAGGAGCT	1014
QY	103	ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu	122
Db	1015	GTTCCTAGGATCGGGGTTGATTGAGAGCTTTCTCGTCACCGCATATCGGTTTTATTG	1074
QY	123	AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle	142
Db	1075	TCAACCCGCACTGCATATGAGCTGCTCCAGAATCGGCAAGGTATTGTCATGGATGTT	1134
QY	143	AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro	162
Db	1135	AATTTCCAGTGAACCAAGCATTCATATCTCTATGAGCAGGGAATCCCTTTGGCTCT	1194
QY	163	LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu	182
Db	1195	CTTTGGGACTTTGGAAAGGCCCAATTTGTTGGAGTCTTTGGTCCACTAGACTTCATCTA	1254
QY	183	IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis	202

Db	1255	ATACTGAGAGAGCTTGGAACTCATGATCAAACTTCACAGAGGAGAGCTTGAGAGCAGC	1314	Unpublished	
Qy	203	ThrIleSerAlaTrpIysGluAlaIysArgGlnThrAsnGlyArgAsnAspSerGlnTrp	222	2 (bases 1 to 1887)	
Db	1315	ACAATAGCAGCTTGGAAAGAGGGAAGGCTCATATTAGCGGCAATATGATGGAATGGG	1374	Buitink, J. and Leprince, O.	
Qy	223	ArgPro---GlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAla	241	Direct Submission	
Db	1375	AGACCATACTTAGGCCACTTGTTGAGTTGGACCTATGATAATCTGAAGAGCTTGCC	1434	Submitted (03-MAR-2003) UMR 1191 Molecular Seed Physiology, INRA,	
Qy	242	ValIysLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAsp	261	16 Bd Lavoisier, Angers 49045, France	
Db	1435	CTGAAATTTTGCAAAACAGGTGGCAGCTGTTCCGGTTATATATCTTCTTTGCCAGAT	1494	Location/Qualifiers	
Qy	262	GlySerPheProGlnLeuHisLeuAlaSerLeuSerGlyIleLeuIysCysIleCys	281	1. .1887	
Db	1495	GGTTCATATCCGAGTTACTGTCATCTTGCTTCGCTATCAGGCATATTAAATGATATGC	1554	/organism="Medicago truncatula"	
Qy	282	ArgTyrPheIysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle	301	/mol_type="mRNA"	
Db	1555	AGATACCTTTAGACATTCGCTAGCTCTTTACCTATCTTCAGCAACCACTTGTCAAT	1614	/cultivar="Paraggio"	
Qy	302	ProLeuGlySerTyrValProIysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu	321	/db_xref="taxon:3880"	
Db	1615	CCCTGGGTACGTGGTCCCTAGATCGGAGATCAAGTAGCAAACTCTCGCTACATTG	1674	/dev_stage="embryo"	
Qy	322	ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer	341	65. .1522	
Db	1675	AGACCTCAGCCCTCTTTGGGTTCTCGCTCGCATATTAGTTCAAGCTGAAGTCAGTTCA	1734	/codon_start=1	
Qy	342	IleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThr	361	/product="AKIN betagamma"	
Db	1735	ATTCCCGTAGTGGATGACACAGCTCGCTTATTTGACATATCTCTCGAAGTATATACT	1794	/protein_id="AA061673.1"	
Qy	362	AlaLeuAlaIysAspIysValTyrThrHisValArgLeuAspGluMetThrIleHisGln	381	/db_xref="GI:29887975"	
Db	1795	GCTCTGGCTAAAGCAAGGCATACGCACAGATTCATCTTGATGACATGACAGTTCACCAG	1854	/translation="MYPSLTDVGGSSSLSSSSSGPILIPKRFVWPGGTRVYLIGSF	
Qy	382	AlaLeuGlnLeuGlyCnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys	401	TRMSEHIPMSPMEGCPVSFQVCSLMPGHQFKFNVDQWRIDEQQEYFVANGYIVNT	
Db	1855	GCCTACAGTTGGGGCAAGATGCGAGCCCGCTTATGGAATCTTCAACGGGCGAGATGT	1914	IYLVREPDILPALISATBSRSHEVDNDFVGAENPRSPSDEVRRIISXFLSE	
Qy	402	GlnMetCysLeuArgSerAspProLeuLeuIysValMetGluArgLeuAlaAsnProGly	421	HTAYDLPSGKVIADVNLPVKQAEHLVYEQDVSMAPLWDFCKSQFVGLSAMDFIL	
Db	1915	CACATGTGCTTGGCTCAGACTCTCTGTGMAAGTGATGGAGCGGTGGCGAATCCAGGT	1974	ILKELGTGSHLIEEQLRTHIAAKGKSKORRALDNNSGNPHCFVHGKESCLKD	
Qy	422	ValArgArgValPheIleValGluAlaGlySerIysArgValGluGlyIleIleSerLeu	441	VALKLVQNKVSTVPIISLEDGSPQLLHLASLGILKACICRHFHSAGSLPILOPIA	
Db	1975	GTAAGGAGGCTGGTGATGTGGAAGCAGGAGCAACGTTGTGAAGGTATCATTTCACTG	2034	SIPLTGTVFNVGDPNGQPLRLRPNASLGALSMFVQAKVSSIPVDENSDLDIYSR	
Qy	442	SerAspIlePheIysPheLeuSerLeu	451	SDITLAKDKAVARIISLDETNIHQALILGQDANSPYGLNNGHRCMCLRSLSLHKWKE	
Db	2035	AGCGATGTTTCCCAATTCCTCGCTCGGTCCT	2064	RLAKPGVRRLVIVEAGSKRVEGIIISLSDVFRFLIG"	
RESULT 9					
LOCUS	AY247268	1887 bp	mRNA	linear	PLN 15-APR-2003
DEFINITION	Medicago truncatula AKIN betagamma mRNA, complete cds.				
ACCESSION	AY247268				
VERSION	AY247268.1	GI:29887974			
KEYWORDS	Medicago truncatula (barrel medic)				
SOURCE	Medicago truncatula				
ORGANISM	Medicago truncatula				
REFERENCE	1 (bases 1 to 1887)				
AUTHORS	Buitink, J., Thomas, M., Gissot, L. and Leprince, O.				
TITLE	Differential expression patterns of beta and gamma subunit isoforms suggest the involvement of different SnRK1 complexes in desiccation tolerance, osmotic stress and starvation in germinating seeds of Medicago truncatula				

Unpublished	2	(bases 1 to 1887)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
-------------	---	-------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 659 CTATGGGATTTTGCAGAGTCAATTTGTTGGAGTCTTAGTGCAATGGACTTCATTC 718

Qy 183 ILeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHis 202

Db 719 ATATTGAAGAGTTAGGAGTCTCATGGATCACTTTGACTGAAGAACAAC-TGAGACTCAT 778

Qy 203 ThrIleSerAlaTrpIysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222

Db 779 ACTATAGCAGCCCTGGAAGAGGGAAATCAAGCAACGGAGGACCTTGATAATAATGAG 838

Qy 223 ArgProGlnGlnHis--LeuValHisAlaThrProTyrGluSerLeuArgAspIleAla 241

Db 839 GGATCAATCTCATTTGTTTTCATGCTGGGCCCAAGGATGCTCTAAAGAGCTGGCT 898

Qy 242 ValIysLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAsp 261

Db 899 TTGAAGGTTTGTCAAAACAAGGTGTCAACTGTTCTCTATCATTT-----TCTTTAGAGGAT 952

Qy 262 GlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCys 281

Db 953 GGTTCATTTCCCAACTGCTGCTGCTTCTTCCATCATGAGAAATTTTAAATGATATGC 1012

Qy 282 ArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301

Db 1013 AGGCATTTGAGCACTCTGCTGGTCTTCTTGCCCATTTCTCAACTTCGGATTGCTTCGATA 1072

Qy 302 ProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu 321

Db 1073 CCTTTGGGTACATGGGTGCTCAAGCTTGGGGATCCAAATGGACACCCACTGATAAGGCTA 1132

Qy 322 ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341

Db 1133 AGCCCAATGCTCTCTGCTGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1192

Qy 342 IleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThr 361

Db 1193 ATACCAATTTGGGATGAAATGATTCATTTCTTGACATATATTCAGAAGTGATATTACT 1252

Qy 362 AlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGln 381

Db 1253 GCATTGGCTAAGGATAAAGCATATGCTCGGATATCCCTGGATGAACTAAATATTTCATCAG 1312

Qy 382 AlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGluArgCys 401

Db 1313 GCATTATTTGGGACAGATGCAATTTCTCTTATGCTTAAACATGGGCACAGATG 1372

Qy 402 GlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGly 421

Db 1373 CATATGTTTGGGCTGCTGATTCATTGACAAAGTGATGGAGCGCTTGGCTAAACCTGGG 1432

Qy 422 ValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeu 441

Db 1433 GTTAGGAGACTTGTGATTTGGAGCTGTAGCAAGCGTGTGGAAGGATATTATTCATTA 1492

Qy 442 SerAspIlePheLysPheLeuLeu 449

Db 1493 AGTGATGTTTCAGATTTCTATTG 1516

RESULT 10

AC128647

LOCUS

DEFINITION Oryza sativa chromosome 3 BAC clone OSUNBb0062G19, complete sequence.

ACCESSION AC128647

VERSION AC128647.8 GI:34447243

KEYWORDS HTG.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzoideae; Oryza.

REFERENCE 1 (bases 1 to 92392)

AUTHORS Buell,R., Hsiao,J., Zismann,V., Moffat,K.M., Hall,J.,

Gansberger,K., Burgess,S., Jarrahi,B., Shvartsbeyn,M., Brenner,M., Ciocko,A., Pai,G., Vanaken,S., Hansen,C., Utterbach,T., Feldblyum,T., Khaliak,H.G., Yuan,Q., Quackenbush,J., White,O., Salzberg,S. and Fraser,C.

Oryza sativa ssp. japonica cv. Nipponbare OSUNBb0062G19 BAC genomic sequence

Unpublished

2 (bases 1 to 92392)

Buell,R.

Direct Submission

Submitted (20-JUL-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

3 (bases 1 to 92392)

Buell,R.

Direct Submission

Submitted (10-APR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

4 (bases 1 to 92392)

Buell,R.

Direct Submission

Submitted (04-SEP-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

On Sep 4, 2003 this sequence version replaced gi:28827864.

Location/Qualifiers

1. 92392

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/sub_species="japonica"

/db_xref="taxon:39947"

/chromosome="3"

/clone="OSUNBb0062G19"

ORIGIN

Alignment Scores:

Pred. No.: 1e-08 Length: 92392

Score: 1171.50 Matches: 389

Percent Similarity: 23.94% Conservative: 34

Best Local Similarity: 22.01% Mismatches: 25

Query Match: 50.02% Indels: 1321

DB: 8 Gaps: 11

US-09-857-525C-2 (1-451) x AC128647 (1-92392)

Qy 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22

Db 76148 GAACATTTGCCAATGTCTCCATGTTGAAGTGTGCCACCGTTTTTCAGGCCATATGCAGT 76267

Qy 23 LeuSerProGlyIleHis----- 28

Db 76208 TTGTCCTCGGATTCATCAGTGGCCCTTTTATCCAAAGTTTATCATAGTTGCTTAG 76267

Qy 28 ----- 28

Db 76268 ACTAGCTTTGCTACCAATAAATTATGTTCTGTGATGCAACATCTAATGTGTGAACATT 76327

Qy 29 -----GluTyrLysPhePheValAspGlyGluTrpArgHisAspG1 42

Db 76328 CTTTAAATCATTTGATATTCAGTACAGTTTGTGTGGATGGGAATGGCGCACGATGA 76387

Qy 42 uArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgG1 62

Db 76388 GCGGCAACCTACCAACAGGAGACTGTGTGGTTAACTTTATGCTTCACTAGGGA 76447

Qy 62 uTyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAs 82

Db 76448 CTTTGACCAGATAAATAATAATTGAGCCCTAGTAGTACCTGGAAGTAGGATGAATGGA 76507

Qy 82 pValAspAsnGlnAsnPheGlnArgThr----- 91

Db 76508 TGTGGACACGACATTTTCAACGTACTGTAAGTGCCTTATCTTGTGTTTGAATTTGA 76567

Qy 91 ----- 91

Db	76568	AAATACTGTAGTAGCTGCTATTATTTCACAGTTCGGAGAAAACAGTGACACTGTGGTA	76627	Db	77647	ATCTTTCAATATCTTCATGA-ACAGGTGACTATGCCTTCTCTCTGCTAGTTTCCCTCAA	77705
QY	91	-----	91	QY	155	-----	155
Db	76628	TACTTGTTCGGCAATTTATCTCTGGTTGATAATAATAGTAGGGTAGTAGTCTCGAA	76687	Db	77706	TCCTTCAATCTTCTCTGCTCTGTAGCTGTTCCCTCCAAATCTTTCATTAGTTGTTCTG	77765
QY	91	-----	91	QY	155	-----	155
Db	76688	GCAATACATGTTGGCAGATACAAATAACACACACTTTAATAGCATTCABAGGTAC	76747	Db	77766	ACATTGAACATAGTAGATATTTTGTCTTGATCTGTAAATATAGTGGAAATAGTACATACCT	77825
QY	91	-----	91	QY	155	-----	155
Db	76748	ACAAATTTGAAATAACATCTAGCCCATCCCATCTGTGAGGCAATTTGATGCTACCT	76807	Db	77826	TTTAGTAAATATGTTGCCCTGTTGTTGGTACAAATGTTTGGTTTCATTGGCAGCTTCAAT	77885
QY	91	-----	91	QY	155	-----	155
Db	76808	GCGAATTTATACATCTTTTCGGTGTACTGAGCAAGTGTGATCAATGAAATGCTGGAGA	76867	Db	77886	TCATCTGAGATTGATGAAATGGAACACATGACGCTCATGCCCTTAAAGTTGACAGTA	77945
QY	91	-----	91	QY	155	-----	155
Db	76868	AGGGGTGAGTTGTAATATACAGCACCTGTGTGCAAAATACAAACGTTCTCTCTCTATT	76927	Db	77946	CTTCACCTACTTCACACTTCTTGTACAAACCAATGTTGCATCCTCTCTCTGTACACATC	78005
QY	91	-----	91	QY	155	-----	155
Db	76928	GGAATAATTTATTTATAGGAAGCGATACAAAGCTGAGTAGCTGATACATTTATTTTATCT	76987	Db	78006	ACCTAGAATCAATGTGTCAACCTTTTAAAGCTTTGAAAGACACTATATATGTAACAA	78065
QY	91	-----	91	QY	155	-----	155
Db	76988	TCATAAAAGTTCAATGTGTGATACATTTATTTTATCTTCATAAAAGTTTATGTG	77047	Db	78066	TTCACTTTAATTTATTAATCGATACATCAATCAATGGGGCTTAAATTTTGTTAAGATTGA	78125
QY	91	-----	91	QY	155	-----	155
Db	77048	TGTGATACATTTCTTTATCTTCACTAABAAAGTTCAGGTGTGTGATACATTTCTTTTA	77107	Db	78126	TTAGAGCAAGTAATTTGTCAAGGGTGTGATTTGATGTCTCAGACATCCAAATACAACTTA	78185
QY	91	-----	91	QY	155	-----	155
Db	77108	CTTCACTAAAGATTCAATGTGTGTGATACATTTATTTTATCTTCACTAACAATTTTCA	77167	Db	78186	GAAGAGTTGGAGAGTAGGTTCTATCTGTTTGCATCCACAATTTCCATATAGTTA	78245
QY	91	-----	91	QY	155	-----	155
Db	77168	GTGTGTGATACATTTCTTTTATCTTCACTAABAAAGTTCATGTGTGTGATACATTTATTT	77227	Db	78246	TAAGAAAGTTCAAGCTCTGTGCTTCTTCTAGCTGCCATCTTGATCTGTGATTTAATTTT	78305
QY	91	-----	91	QY	155	-----	155
Db	77228	TTATCTTCACTAAAGTTTGTGTGTGCGCAATACATTTCTTTTATCTTCACTAACAAGT	77287	Db	78306	GGCATCTTGATATAAAGTGAATTTTCATAGTTTCTTTAACCATATATTAAGTACAAG	78365
QY	91	-----	91	QY	156	-----	159
Db	77288	CATGTGTGTTTATAATGCTTCTTCTTCTTGTGTTAGTTTGGCTTTTGTCTTTTCACTTTC	77347	Db	78366	AAATATTGGCTGTTTGGGAACATACATTTAAATAAATAGTGTACCTTGCAGGGAATCC	78425
QY	92	-----	94	QY	159	oValAlaProLeuTirAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAs	179
Db	77348	TTTCTAGAGCTAAATATAATAAACAATTTTCTTGTGTTTCTTCAAGCAGGTTCTTT	77407	Db	78426	TGTGGCACCTCTATGGGATTCATTCAGGGGCCAGTTTGTGGCTTTTGAGCCACATGA	78485
QY	94	sSerAspGlyThrValSer-GluGlyThrLeuArgValSerGluAlaAlaIleGlnIleSe	114	QY	179	pPheIleLeuIleLeuArgGlu	186
Db	77408	GTCTGATGCAATTAATTCAGGAAGGCTCTCAGAGAAATTCAGAGCAGCTATACAGATCTC	77467	Db	78486	TTTTTATCTCATATTGAGAGAGGTAATTTATCTGTCTCACTTACTGTTTGGTGTAGTATT	78545
QY	114	rArgCysArgValSerGluTyrlLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSe	134	QY	186	-----	186
Db	77468	TAGGTCTGTGTAGCGGATTTCTGAATGAGCAAACTGGGTATGATTTACTCCAGATTC	77527	Db	78546	CTTTAATTTATTTTATGTGTCATTTGTTTCTCTATTGTATTTTATTATTATTATTACAA	78605
QY	134	rGlyLysVal	137	QY	186	-----	186
Db	77528	TSGCAAGGT-ATGGAATAGCTCGTCTGTATGAATTTTACTGCTCTAAATGTTATTTC	77586	Db	78606	GTGATGGTTTAAATGTTTCAATTCAAATATGCTGCCACTTTTAGTGGCCCTGCAAAAAGAA	78665
QY	138	-----	148	QY	186	-----	186
Db	77587	TTATTACATTTATATTACAAATTCAGGTCAATGTCTCTAGACGTTAAATTTGCTGTGAAGA	77646	Db	78666	TTGCGCAATCTGTGTTTTTTTTTTTTTTTTTTTTTAAATAAATCTGAAAGCAGTGAATATCATCTTTG	78725
QY	148	nSerPheHisIleLeuHisGlu	155	QY	186	-----	186
				Db	78726	ACAGTATTAGTTATTGACCTAGAGCTTATTTCACATGAATTTATAAGACGACTGGTTC	78785

Db 80946 GGCTGAGTTTGATGTGTTTGGTTCATTTATCACTCTTACTTCTCCATACCGGTCAACAA 81005

Qy 380 ----- 380

Db 81006 ACATATACATATAAGATAAGTAGTCAATGCTGTTTAAAGCAAGCTTCTCGTGAATT 81065

Qy 380 ----- 380

Db 81066 TGTGTATATGTTATCCATCGAGCTGGTGTGATTTTGGTGTGAACCTACTCATCTTTTGT 81125

Qy 381 -----GlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAspG 398

Db 81126 GTTATGTTTCAGGCTCTGCAGCTTGGACAGCATGCAAAATTCGCCCTTTGGATTCTTTAACG 81185

Qy 398 LysGlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuA 418

Db 81186 GGCAAAGATGTCAGATGTGTCCTCGTCTGACACTTTGTTGAAGTAGTATGAGCGATTAG 81245

Qy 418 laAsnPro----- 420

Db 81246 CTATCTGGTAATTATCTGTGCTAGATATGCTTCTGAAGCAACTTTTCGGGTCAA 81305

Qy 421 -----GlyValArgArgValPheIle 427

Db 81306 CTTACTAATATTTTCTTTTTCATTTTATATGTTGCTAGGGTGGCGCTGTTTCATT 81365

Qy 428 ValGluAlaCysSerLysArgValGluGlyIleLeuSerLeuSerAspIlePheLysPhe 447

Db 81366 GTGGAGCTGGTAGCAACACGTGTGGAGGGCAATAATATCACTGAGTGATATTTTCAAGTTC 81425

Qy 448 LeuLeuSer 450

Db 81426 TTGCTGAGC 81434

RESULT 11

AC096690/c

LOCUS

DEFINITION

Oryza sativa chromosome 3 BAC OSJNBa0059G06 genomic sequence,

complete sequence.

AC096690

AC096690.5 GI:29568036

HTG.

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Eleutheroidae; Oryzaceae; Oryza.

Buell, C.R., Yuan, Q., Qiyang, S., Liu, J., Gansberger, K., Jones, K.M.,

Overton II, L.L., Tsitrin, T., Kim, M.M., Bera, J.J., Jin, S.S.,

Padrosh, D.W., Tallon, L.J., Koo, H., Zismann, V., Hsiao, J., Blunt, S.,

Vanaken, S.S., Riedmuller, S.B., Utterback, T.T., Feldblyum, T.V.,

Yang, Q.Q., Haas, B.J., Suh, B.B., Peterson, J.J., Quackenbush, J.,

White, O., Salzberg, S.L. and Fraser, C.M.

Oryza sativa chromosome 3 BAC OSJNBa0059G06 genomic sequence

Unpublished

REFERENCE

1 (bases 1 to 145115)

Buell, R.

Direct Submission

Submitted (22-SEP-2001) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA

3 (bases 1 to 145115)

Buell, R.

Direct Submission

Submitted (08-MAY-2002) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA

4 (bases 1 to 145115)

Buell, R.

Direct Submission

Submitted (05-APR-2003) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA

5 (bases 1 to 145115)

AUTHORS TITLE JOURNAL COMMENT

Buell, R.

Direct Submission

Submitted (25-APR-2003) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org

On Apr 5, 2003 this sequence version replaced gi:20503077.

Address all correspondence to:rice@tigr.org

BAC clone OSJNBa0059G06 is from Oryza sativa chromosome 3

The orientation of the sequence is from SP6 to T7 end of the BAC

clone.

Genes were identified by a combination of several methods: Gene

prediction programs including Fgenesh (<http://www.softberry.com/>),

genscan and Genscan+ (Chris Burge,

<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (Mark Borodovsky,

<http://genemark.biology.gatech.edu/GeneMark/>), and GeneSplicer

(Mihaila Pertea and Steven Salzberg, contact mpertea@tigr.org),

searches of the complete sequence against a peptide database and

the plant EST database at TIGR (<http://www.tigr.org/tdb/tga.shtml>).

Annotated genes are named to indicate the level of evidence for

their annotation. Genes with similarity to other proteins are named

after the database hits. Genes without significant peptide

similarity but with EST similarity are named as unknown proteins.

Genes without protein or EST similarity, that are predicted by more

than two gene prediction programs over most of their length are

annotated as hypothetical proteins. Genes encoding tRNAs are

predicted by tRNAscan-SE (Sean Eddy,

<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are

identified by RepeatMasker (Arian Smit,

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

This BAC overlaps with rice BAC OSJNBb0062G19 (AC128647) and

OSJNBa0033P04 (AC092263).

FEATURES SOURCE

1. 145115

Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="3"

/map="near L375"

/clone="OSJNBa0059G06"

/note="japonica cultivar-group"

1271..1331

/rpt_family="AT-rich"

2342..3105

/gene="OSJNBa0059G06.1"

/note="predicted by genemarkHM"

join(2342..2619,2999..3105)

/gene="OSJNBa0059G06.1"

join(<2342..2619,2999..3105)

/gene="OSJNBa0059G06.1"

/codon_start=2

/product="hypothetical protein, 5'-partial"

/protein_id="AAP21395.1"

/db_xref="GI:30102982"

/translation="EQMRRAAGGGHRCRCGRWRARRGQLRGVEAAAGGRRRG

RGGLCRHRTAAASSHSIHVARELPPELSLNLEFLTSSWWSASAVDFQASGAVNDSE

RLLRHPAAEAANTARGASASTVGR"

2353..2479

/rpt_family="G-rich"

3392..3467

/rpt_family="(CCG)n"

5230..5343

/rpt_family="(TATAA)n"

5868..8753

/gene="OSJNBa0059G06.2"

/note="similar to family II lipase EXL3 GB:AAK30018

GI:15054386 (Arabidopsis thaliana); contains Pfam profile

PF00657 (GDSL like Lipase Acylhydrolase)"

join(<5868..6150,7939..8753)

/gene="OSJNBa0059G06.2"

join(5868..6150,7939..8753)

/gene="OSJNBa0059G06.2"

repeat_region

gene

mRNA

CDS

repeat_region

repeat_region

repeat_region

gene

tRNA

CDS


```
/codon_start=1
/product="putative lipase/acylhydrolase"
/protein_id="AAP21383.1"
/db_xref="GI:30102970"
/translation="MTMKKQKQNMKLGILLVAVAVAMAPAGPASFSSGRGAPMPAV
IVGDSIVPGNNNKTKQIKAHAPYGMDFANSEPTGRVSNGLIPTDFIVQGLNVKQ
LMPPTGLVSLPDLTKGSPASGATGYDPLTVIVSVLTLXQIQIYEVEHYRKLGV
VGEETARILGVALFVVCAGTDIDANTFYFTTPRSVEYDIPSYVDLLVSAKLLQV
AALGARIRIGVGLPIPGVPSQRTILGGPHRRCEERENYAAKLFNSRMEVIAAKNP
ATRMVVDIYITLQELVENDKYGFETTRGCCGTTEVTGLCDARFVDICDNVSN
HVFDSVHTQRAYKIIVDYIFDNYLQFLIA"
6410..6437
/rpt_family="AT-rich"
/781..7804
/rpt_family="(CGAT)n"
12122..13806
/notes="OSJNBa0059G06.3"
/notes="predicted by fgenesh"
join(12122..12362,12851..13020,13702..>13806)
/genes="OSJNBa0059G06.3"
/notes="OSJNBa0059G06.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP21385.1"
/db_xref="GI:30102972"
/translation="MRVRFKVGASPGFKESVSGACWNGAPRZAGDERPPGADGNG
SEAMPIDGGARKWFGVIGSSRTVFARVVDLGGDLDARKNMAMWMAKAAVDQTAEP
SSRCGLARGSESQPCSELAAMRQSMVTVAEVELDSGNREESRDGFKVDSIRGIG
TSIWRFPKGD"
13321..13364
/rpt_family="CA-rich"
15008..15049
/rpt_family="(TC)n"
16325..16352
/rpt_family="(TTCC)n"
16964..17005
/notes="overlap BAC OSJNBa0033P04 has 49 TA repeats in this
region."
16964..17005
/rpt_family="(TA)n"
17177..17203
/rpt_family="AT-rich"
17384..18823
/genes="OSJNBa0059G06.4"
/notes="contains Pfam profile PF01554 (MatE)"
<17384..>18823
/genes="OSJNBa0059G06.4"
17384..18823
/codon_start=1
/product="putative MatE efflux family protein"
/protein_id="AAP21388.1"
/db_xref="GI:30102975"
/translation="WTPPPPGSPHERKTKWAESVASEFRAQGIAPPLIAMNLTWFAKL
AVTAFLEGIDQLAAGTGLGFSFANVTGFVILGLCAAMDPICGOAGHANGKILRK
TLVMTATLIGASIPATLMLHVDALLRFQOQADMSNARSVYVCLLPDLAVTSFVN
PLKXSLQRQGTPLTPASALALAHVPLTWMTTRGIQGVATVAVVSDLAVALMLA
GVLVSRKRRKAGGGGWEQTRGRENRLRLAVPSCNLCLEWVCYVLLVLTGRLP
DARTVAVMTNFDYLLFAGMLSLVSASVRVNEGLGAGHMAARAGMVSIVGGA
VGGVGGVAMVAAARAWGSIYSDAGVRGVGRVMEVAVLEVVNPFNVCGIVRGTT
ARPVGVMVAVAGFVYVLLPLGLVALAFKARLGIQGLLGLVYGAASLAVLTFIARM
DKPAEQAKRTTATVAQFHQHVWQV"
18325..18515
/rpt_family="G-rich"
20002..21960
/genes="OSJNBa0059G06.5"
/notes="EST CA759271.BE340176 from this gene"
join(<20002..20123,20242..20298,20375..20530,20579..20654,
20811..20903,21132..21245,21366..21960)
/genes="OSJNBa0059G06.5"
join(20002..20123,20242..20298,20375..20500,20579..20654,
20811..20903,21132..21245,21366..21419)
```

```
/gene="OSJNBa0059G06.5"
/codon_start=1
/product="unknown protein"
/protein_id="AAP21391.1"
/db_xref="GI:30102978"
/translation="MAPPPCCVUVRCCSGSLPSPPPPPPPSPPSNNLRARREVF
LRSELAATLAAIFHLGSGTKPRYLGVSQPSLALCPAINNCVSTCEIDTISIHVAPW
NYPKQORRAKPIITKHEAINQLIQVVTQTPDNFTPLRVEKTDYDVRVEYIFGFV
DDVEFFPPCKNSIVQVSASRSGFIDFNANKKRVKALRLALENKGWASBSTI"
20063..20084
/rpt_family="(CCG)n"
20673..20703
/rpt_family="(TTTT)n"
22163..22193
/rpt_family="AT-rich"
24255..25676
/genes="OSJNBa0059G06.6"
/notes="predicted by fgenesh"

Alignment Scores:
Pred. No.: 1.81e-88 Length: 145115
Score: 1171.50 Matches: 389
Percent Similarity: 23.94% Conservative: 34
Best Local Similarity: 22.01% Mismatches: 25
Query Match: 50.02% Indels: 1321
DB: 8 Gaps: 11

US-09-857-525C-2 (1-451) x AC096690 (1-145115)
QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
Db 138916 GAACATTGGCAATGCTCTCTGTTGAAGGTTGTCCACCGTTTTTCAGGCCATATGCAGT 138957
QY 23 LeuSerProGlyIleHis----- 28
Db 138856 TTGTCCTCCCTGGGATCACTAGGTGCCCTTTTATCCCAAGTTTATCATAGTGTCTTAG 138797
QY 28 ----- 28
Db 138796 ACTAGCTTTGCTACCCATAAATATATGTTTGTGTGATGCAACATCTAATGTGTGAACATT 138737
QY 29 -----GluTyrLysPhePheValAspGlyGluTyrPArgHisAspG1 42
Db 138736 CTTTAAATCATGTATATTCAGTACAGTTTGTGTGGATGGGGAATGCGCACGATGA 138677
QY 42 uArginProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgG1 62
Db 138676 GCGGCAACTACCATACAGGAGACATGCTGTGTGTTAACTTTATGCTTACTAGGGA 138617
QY 62 uTyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAs 82
Db 138616 CTTTGACCAGATAAATAACAATATTGAGCCCTAGTACACTTGAAGATGAGATGACATGGA 138557
QY 82 pValAspAsnGluAsnPheGlnArgThr----- 91
Db 138556 TGTGACCAACGACAAATTTTCAACGTAAGTGCCACTTATCTGTTTGTAAATTTGA 138497
QY 91 ----- 91
Db 138496 AAATACTTGATAGTAGTCTGCTATTTTTCACAGTTTCGAGCAACACAGTGCACACTGTGGTA 138437
QY 91 ----- 91
Db 138436 TACTTGTGCGCAATTTATCTCTGGGTTGATAATATGATAGGGGTAGTATGCTTCTGAA 138377
QY 91 ----- 91
Db 138376 GCATACCATTGTTGGCGCAGATACATAATACACACACTTTTAATAGCATTACAAGGTCCAC 138317
QY 91 ----- 91
Db 138316 ACAAAATTGGAAATAACATCTAGTCCCATCTGTCTGTCAGCGCAATTTGTATGTGTACCT 138257
```

```
QY 91 ----- 91
Db 138256 GCGAATTTATACATCTTTTCGGTGTACTAGCAAGTGTGATCATTTGAATGCTGGGAGA 138197
QY 91 ----- 91
Db 138196 AGGGTGGAGTGTAAATTATCAGCACCTGTGTGCAAAATACAAATGTTCTCTCCTCTATT 138137
QY 91 ----- 91
Db 138136 GGAATAATTATTATAGAAAGGATACAAGCTGAGTAGCTGTACATTTATTTTATCT 138077
QY 91 ----- 91
Db 138076 TCACATAAAAGTTTCATGTGTGTGATACATTTATTTTATCTTTCATAAAGTTCATGTG 138017
QY 91 ----- 91
Db 138016 TGTGATACATTTCTTTTATCTTCACTAAAGATTCACGTGTGTGATACATTTCTTTTATA 137957
QY 91 ----- 91
Db 137956 TCTTCAATAAAGTTCATGTGTGTGATACATTTATTTTATCTTCACTACAATGTTCAT 137897
QY 91 ----- 91
Db 137896 GTGTGTGACATTTCTTTTATCTTCACTAAAGATTCACGTGTGTGATACATTTTATT 137837
QY 91 ----- 91
Db 137836 TTATCTTCACTAAAGTTCGTGTGTGCAATACATTTCTTTTATCTTCACTACAAGGT 137777
QY 91 ----- 91
Db 137776 CATGTGTGTTTAAATGCTTCTTATTTGTAGTTTTCGCTTTTGCCTTTTACCTTTC 137717
QY 92 ----- ValThrLe 94
Db 137716 TTCTTAGAGCTAAATATAATATAACAACATTTCTTGTCTTCTTCAAGCAGGTTTCTTT 137657
QY 94 nSerAspGlyThrValSerGluGlyThrLeuArgValSerGluAlaLeuGlnIleSe 114
Db 137656 GTCGTGAGCATTTATTCAGGAAGTCTCTCAGAAATTCAGAGCAGCATACAGATCTC 137597
QY 114 rArgCysA-gValSerGluThrLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSe 134
Db 137596 TAGTGTGCTGTAGCGAATTTCTGAATGCACAACTGGGTATGATTTATCCAGATTC 137537
QY 134 rGlyLysVal ----- 137
Db 137536 TGSCAAGGT-ATGGAAAATGAGCTCGTCTTGTATGAATTTTAGTGCTCTAATGGTTATTC 137478
QY 138 ----- IleAlaLeuAspIleAsnLeuProValLysG1 148
Db 137477 TTATTACATTAATTACAAATTCAGGTCAATGTCTAGACGTTAAATTTGCTGTGAAGCA 137418
QY 148 nSerPheHisIleLeuHisGlu ----- 155
Db 137417 AUCTTTTCAATTTCTTCATGA-ACAGGTGACTATGCCCTTCTCTGCTAGTTTCCCTCCAA 137359
QY 155 ----- 155
Db 137358 TCTTTCATTTCTTCTCTGCTCTGTAGCCTGTTCCCTCCAAATCTTTCATTTAGTTGTTCTG 137299
QY 155 ----- 155
Db 137298 ACATTGAACTAGTGATATTTTGTGCTGATCTGTAAATATAGTGGACAGGTACATACCTT 137239
QY 155 ----- 155
Db 137238 TTTAGTAAATATGTTGGCCTGTTGTTGGTGATACATGTTTGGTTTCATTTGCCAGCTTCAAT 137179
QY 155 ----- 155
```

```
Db 137178 TCATCTGAGATTGCATGAATAATGGAAACACTGCAGCCTCATGCCCTTAAAGTTTGAGCTA 137119
QY 155 ----- 155
Db 137118 CTTACCTACTCTTACACTTTCTTGTACAAACCAATGTTGCATCCTTCTCTTGTACACATC 137059
QY 155 ----- 155
Db 137058 ACCTAGAAATGCATGTGGTCAACCTTTTAAAGCTTGAAAGAGCACATATATATGTAAACACA 136999
QY 155 ----- 155
Db 136998 TTCACTTTAAATTATTAATCGATACATCATCAATGGGGCTTAATTTTGTTHAGATTGCA 136939
QY 155 ----- 155
Db 136938 TTAGAGCAAGTAATTGTCAAGGTGATGATGTGTCAAGACATCCAAATACAACAAATTA 136879
QY 155 ----- 155
Db 136878 GAAGAGGTGGAGGAAGTAGGTTCTATTCCTGTTGCACTCCACAATTTCCATATAGGTTA 136819
QY 155 ----- 155
Db 136818 TAAGAAAGTTCAGCTCTGTGCTTCTTAGCTGCCATCTTGATTCGTGATTTAATTAATTTT 136759
QY 155 ----- 155
Db 136758 GGCATCTTGATATAACTGGAATTTTCATAGTTTCTTTAAACCATATATATTGAAAGTACAAG 136699
QY 156 ----- GlnGlyIlePr 159
Db 136698 AAATATTGGCTGGTTTGGGNAACATACATTAATAATAGTGTACCTGCGAGGGAATTC 136639
QY 159 oValAlaProIleuTrpAspSerPheA-gGlyGlnPheValGlyLeuLeuSerProLeuAs 179
Db 136638 TGTGGCACCTCTATGGGATTCATTCAGGGGCCAGTTTGTGGCCCTTTTGGACCCCACTGGA 136579
QY 179 pPheIleLeuIleLeuArgGlu ----- 186
Db 136578 TTTTATACTCATATTAGAGAGGTAATTTATTCGTCACTTACTGTTTGGTGTAGTATT 136519
QY 186 ----- 186
Db 136518 CTTTAAATTTTTTTATGTCATGTTGTTCTCTATTTGTATATTTTTTATTTGTTACAA 136459
QY 186 ----- 186
Db 136458 GTGATGGTTTAAATGTTTCATTTCAAATATGCTGCCACTTTTAGTGCCCTGCAAAACGAAA 136399
QY 186 ----- 186
Db 136398 TTGCGCAATCTTGTGTTTTTTTTTTTTTAAATAATCTGAAGCAGTGAATATCATTTCTTTG 136339
QY 186 ----- 186
Db 136338 ACAGTATTAAATTATTTGACCTAGAACCTTATTTACATGAATATATAAGACGACTGGTTC 136279
QY 186 ----- 186
Db 136278 ATTAATGGAATGGGAGTCTATGCATTTGAACAACTGACCTGCTTCTGTAATCCGATTC 136219
QY 186 ----- 186
Db 136218 GCTGTATCACTTAGTAACCATGCATCTCATGGAACCTTTGTTGGGTGACCACAGAAGTT 136159
QY 186 ----- 186
Db 136158 CAGCCAAGATAGTGAATCATGTTGAACCTGTTTCCCTTAATTTGGAAGTGTTCAGTTAGTA 136099
QY 186 ----- 186
```

Db 136098 GAACCTCCAAAATTTGAAATCGTATCTTTTGGAGGTGGACTCTAGAAAATAGAAATRA 136039
 Qy 186 ----- 186
 Db 136038 ATATGCTAGTAGAGCGATGCATCATGATACCTTTATTTATTTATGCTGCTTGAATGCTTC 135979
 Qy 186 ----- 186
 Db 135978 TGCTTCCAGGTAAAGGATTAATCTCCAGCATGTGATTTGTACTGTGAACAATTTGT 135919
 Qy 187 ----- LeuGluThrHisG 191
 Db 135918 GAGCATATTCACACAGTAGCATTAATATCTTGTTATTTATTCGAGCTGGAACTCATG 135859
 Qy 191 lySerAsnLeuThrGluGlnLeuGluThrHisThrIleSerAlaTrpLysGluAlaL 211
 Db 135858 SCTCCAACTGACAGAGAGCAGCTTGAACACATACATATCTGCTGGAAGAGGCCA 135799
 Qy 211 yArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHisLeuValHis- 230
 Db 135798 AGCGTCAAACTTATGCCAGAAACGAGGGTTCTTGGAGGGCAAAATCACCAITTAGTTCATG 135739
 Qy 230 ----- 230
 Db 135738 TGAGTTGTCCACTGTGTTTGGAAACCAATTGCARAATCTGTATCGCACTCACCAAAAATTGA 135679
 Qy 230 ----- 230
 Db 135678 GACTGTGCCATCAACTTCTATCTATATTTACTGGCCAAATTTACTCCAGTGAATTTTA 135619
 Qy 230 ----- 230
 Db 135618 GAATTCAGTATTCATTCAGGATTAATATTTTACCTTTGATTTATTTGCAATTCACA 135559
 Qy 230 ----- 230
 Db 135558 AATTGTTGCACAACTGTTGTGTGAGTACCACATAACAGTTTTTGGCACACCATGACAC 135499
 Qy 230 ----- 230
 Db 135498 ACATTTTGTATACAACTTTTGGCACATGATGACACATTTTACTTTTGTGTACAAC 135439
 Qy 230 ----- 230
 Db 135438 TAAGTTACTACTGATGCTGAATATATTTTGTCTTAATTTAGCTTTAATTCCTGCTCA 135379
 Qy 230 ----- 230
 Db 135378 AGTGAGTTACTTAGTGAACTGCAATTTGTTTGTGTGAAATTTTCATCGAACCCCTTGCA 135319
 Qy 230 ----- 230
 Db 135318 CAAGTGTCTTTTCAATTTCCATCCTAGTATTTTATTAACATGCTTAGTTTATATGATC 135259
 Qy 231 ----- AlaThrProTyrGluSer 236
 Db 135258 ATGTGCTTCATGCACTATGCTGTGAGTGAATCTATTTGTTAGGCCACCCCTTATGAATCC 135199
 Qy 237 LeuArgAspIleAlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyr 256
 Db 135198 CTGAGGGAAATTTGCTATGAGATACCTGCAAAACGGTGTCTTACCGTTTCCAAATATATGTT 135139
 Qy 257 SerSerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIle 276
 Db 135138 TCCTCATCACAGATGGCTCATATCCAAATGTTGCAATCTGTCTTCCCTTCTCGGAAT 135079
 Qy 277 LeuLysCys ----- 279
 Db 135078 TTGAAATGTAAGTGTCTTATGTTATTTGCAAGAAAGTGTCTATACAAACATATTTAC 135019
 Qy 280 ----- IleCysArgTyrPheLysAsnSerThrG 289
 Db 135018 ATGATAGCTTAACAACATTTTCATTTTTCAGGATTTTGTAGATATTTTAAATAATCTCTCAAG 134959

Qy 289 yAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerIleProValProLy 309
 Db 134958 TAATTTACCTATTTTGGCCACCTGTATGCAATTCCTCTGGTACCTGGTTCAAA 134899
 Qy 309 sIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSerSe 329
 Db 134898 AATTGGTGATCCTAATAGGCGTCCATTGGCTATGTTGGGCGCTTAACACATCTCTTAGCGC 134839
 Qy 329 rAlaLeuAsnMetLeuValGln ----- 336
 Db 134838 TGCCCTGAATTTGCTGCTGTTCAAGGTATGTTATCTGCTGTTTACCATAACCAATGCGCTG 134779
 Qy 336 ----- 336
 Db 134778 GGAAGTACTTGCATTTCTGTTCTTGGTGCATTTACTGAAGTTTAATACTAATTCCTACT 134719
 Qy 336 ----- 336
 Db 134718 GAGATCTCTTACAGCTTTTCGATTATTCGCTATCTTGTGTCCTATTAATGTCTTTATTA 134659
 Qy 337 ----- AlaGlyValSerSerIleProIleValAs 346
 Db 134658 CAATTACTCTGTACTCTAATTTGTTTCAATCAGCTGGTGTGAGCTCAATACCAATTTGGA 134599
 Qy 346 pAspAsnAspSerLeuLeuAspIleThrIleValSerArg ----- 357
 Db 134598 TGACATGATCTCACTGCTTGCACATATTCAGAGAGTGCCTATTCGATACTACCATTTC 134539
 Qy 358 ----- S 358
 Db 134538 TTTGTGTAGTCATATACATATTTCTTTGTGCTCATTTACTTGTGCTTTTCCATATGTA 134479
 Qy 358 erAspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetT 378
 Db 134478 GTGACATCAACAGCTCTTGCARAAGACAGGTGTACACACATTCGCTAGATGAGTGA 134419
 Qy 378 hrIleHis ----- 380
 Db 134418 CCATTCATCAGGTATGCTAGTAAATATTTAATGCTCTTGTATTTAGTTTGTCTGGAAT 134359
 Qy 380 ----- 380
 Db 134358 TTTAGCTGGGATACAGTTGTACATAATTTTTCAGAGATTTCTTTTGGTCTTGAGACTG 134299
 Qy 380 ----- 380
 Db 134298 GTTAACTTAGAAATAATAACTTTTACTGATTTATGATATAGAGGAAATCTGTATCGA 134239
 Qy 380 ----- 380
 Db 134238 TACCTGCATCACATATATATATGAGGACATGCTTTGTGTTCTGGCAATTTATAGC 134179
 Qy 380 ----- 380
 Db 134178 CCTGTGCTTTTTCACGACAGAGCATGAGCATGCACACATCTAAACAAAGGCTGTA 134119
 Qy 380 ----- 380
 Db 134118 GGCTGTAGTTGATGTGTTTGTTCCTATTTATCACTCTTACCTTTCCATACCGGTCAAACA 134059
 Qy 380 ----- 380
 Db 134058 ACATATACTATAATAGATAAGTAGTGAATGCTGTTTAAAGCAGGCTTTCTCTGTGAAT 133999
 Qy 380 ----- 380
 Db 133998 TGTGTATATGTTATCCATTCGAGCTGGTGTGATTTTGTGTGAACCTACTCATTTTGT 133939
 Qy 381 ----- GlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnG 398
 Db 133938 GTTATGTTTCAGGCTCTGAGCTTGGACAGGATGCAAAATTCGCCCTTTTGGATTTCTTAACG 133879

QY 398 lyGlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuA 418
 Db 133878 GGCAGAGATGTCAGATGTCCTCCGCTCGACACTTTGTGTGAAGTGTGGAGCGATTAG 133819
 QY 418 laAsnPro----- 420
 Db 133818 CTAATCCGTGTAATTAATCTGTGCTAGATATGCTTCTCGAAGCAACTTTTCGGGTCAA 133759
 QY 421 -----GlyValArgValPheIle 427
 Db 133758 CTTACTAATAATTTTTTTTGGCAATTTATCTGTGTAGGGGTGGCCGCTGTTTTCATT 133699
 QY 428 ValGluAlaGlySerLysArgValGluGlyIleLeuSerLeuSerAspIlePheLysPhe 447
 Db 133698 GTGGAAGCTGGTAGCAACGCTGTGGAGGCTAATATCACTGAGTGATATTTTCAAGTTC 133639
 QY 448 LeuLeuSer 450
 Db 133638 TTGCTGAGC 133630
 RESULT 12
 LOCUS BT009623 982 bp mRNA linear PLN 20-JUN-2003
 DEFINITION Triticum aestivum clone wreln.pk0143.e2.fis, full insert mRNA
 sequence.
 ACCESSION BT009623
 VERSION BT009623.1 GI:32129174
 KEYWORDS FLI CDNA.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticaceae; Triticum.
 1 (bases 1 to 982)
 Tingey,S.V., Walters,P., Powell,W., Dolan,M., Miao,G.-H.,
 Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
 Direct Submission
 Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and
 Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
 USA
 FEATURES
 source Location/Qualifiers
 1..982
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="wreln.pk0143.e2.fis"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,02e-82 Length: 982
 Score: 1068.00 Matches: 205
 Percent Similarity: 95.61% Conservative: 13
 Best Local Similarity: 89.91% Mismatches: 10
 Query Match: 45.60% Indels: 0
 Gaps: 0
 Db: 8
 US-09-857-525C-2 (1-451) x BT009623 (1-982)
 QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrShuSerLeuArgAspIleAlaVal 242
 Db 11 AGATCAATCAGCACTAGTATGATGTCACCCCTTATGAATCCTTGAGGGGTATTGCCATG 70
 QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAspGly 262
 Db 71 AAAATACCTGAACTGGCATTTCTACAGTCCCAATCATCTATTTCATCGTCATCAGATGA 130
 QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
 Db 131 TCGTTTCCGAGCTGTGTGATCTGTGATCCCTTTCAGGAATTTTGAAATGATCTCTGAGA 190
 QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
 Db 191 TACTTCAAGAACTCCACTGGTAGTTTGGCGGATTTCTAAACCAACCATGATCTCAATTCOG 250

QY 303 LeuGlySerTyrValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
 Db 251 CTGGTACCTGGGTTCAAAATTTGGTGAACCAATGGTCAATCCATTGGCTATGTCGG 310
 QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
 Db 311 CCTAATACATCTCTTAGCTCTGCTTAACTTGTGTGGTCAAGCTGGAGTTAGTTCATA 370
 QY 343 ProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
 Db 371 CCCATTGTGATGATACAGACTCGCTGATGACACATCTCCACAGAGTGCATCACAGCT 430
 QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
 Db 431 CTACGGAAGACAGTCTACACCCATATCCGCTAGATGAGATGACCATTCATCAGGCC 490
 QY 383 LeuGlnLeuGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
 Db 491 TTGCAGCTCGGGCAAGACGCGAATTCACCTTTTGGACTTTTCAATGGTCAAGATGCCAG 550
 QY 403 MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
 Db 551 ATGTGTCCTCAGCTCTGACCCCTTTGCTGAGGTTATGGAGATTTGGCTAATCTCGGGTG 610
 QY 423 ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleLeuSerLeuSer 442
 Db 611 CGTCCGCTGTTTCATCGTGAGCTGGCAGCAAGCGAGTGAAGGCGTAAATATTCGCTGAGC 670
 QY 443 AspIlePheLysPheLeuLeuSer 450
 Db 671 GACATATTAAGTTGCTGCTGAGC 694
 RESULT 13
 LOCUS F7G19 110514 bp DNA linear PLN 16-SEP-1997
 DEFINITION Sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1,
 complete sequence.
 ACCESSION AC000106
 VERSION AC000106.1 GI:2342673
 KEYWORDS HTG.
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 110514)
 Osborne,B.I., Vysotskaia,V.S., Toriumi,M., Yu,G., Oji,O.,
 Buehler,E., Conway,A.B., Conway,A.R., Dewar,K., Feng,J., Kim,C.,
 Kurtz,D., Li,Y., Shinn,P., Sun,H., Davis,R.W., Ecker,J.R.,
 Federspiel,N.A. and Theologis,A.
 The sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1
 Unpublished (1997)
 2 (bases 1 to 110514)
 Theologis,A.
 Direct Submission
 Submitted (20-JAN-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 3 (bases 1 to 110514)
 Theologis,A.
 Direct Submission
 Submitted (17-FEB-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 4 (bases 1 to 110514)
 Theologis,A.
 Direct Submission
 Submitted (27-FEB-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 5 (bases 1 to 110514)
 Theologis,A.
 Direct Submission
 Submitted (03-MAR-1997) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA

REFERENCE AUTHORS TITLE JOURNAL	6 (bases 1 to 210514) Theologis,A. Direct Submission Submitted (02-Apr-1997) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA 7 (bases 1 to 110514) Theologis,A. Direct Submission Submitted (16-Apr-1997) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA 8 (bases 1 to 110514) Theologis,A. Direct Submission Submitted (26-Aug-1997) Plant Gene Expression Center, 800 Buchanan St., Albany, CA 94710, USA 9 (bases 1 to 110514) Theologis,A. Direct Submission Submitted (16-Sep-1997) Plant Gene Expression Center, 800 Buchanan St., Albany, CA 94710, USA On Aug 26, 1997 this sequence version replaced gi.1841948. The sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1. Location/Qualifiers 1. 110514 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /culivar="Columbia" /db_xref="taxon:3702" /chromosome="1" /clores="F7G19" 971..2850 /gene="F7G19.1" join(971..1279,1540..1641,1688..1837,2134..2232, 2556..2665,2740..2850) /gene="F7G19.1" /note="Similar to ATP-dependent Clp protease (gb D90915). EST gb M65461 comes from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAB70396.1" /db_xref="GI:2342674"	gene
CDS	/translation="MASCQASMSLPRSSFPHPLSSNSGRRLNKTFRYAFA KASAKEMPPINPDPFLSTLAASPEKLNRPNADVPYDIDPDSQMSSPA QVSRYAHERPTPPDPSMLDGRIVIGMTDDDDIFCQLVPAVTLVVAELM YLOWLDEKPIYIVINSTGTRDDGTWHTVCVGAACQACLLISAGTKGRFMEHA KGESEFDYFELIPVMQLQVITNRDILVELLSKHTGNSVETVANVRRPYMDAKA KEFGVDRVSLGTVF" complement(3700..3771) /product="CRNA-Glu" /note="Codon recognized: GAA" 3942..6513 /gene="F7G19.2" join(3942..4046,4391..4507,4747..4854,4973..5164, 5263..5454,5551..5669,5747..5834,5916..6035,6216..6356, 6436..6513) /gene="F7G19.2" /note="Similar to probable Mg-dependent ATPase (pir S56671). ESTs gb T46782,gb AA:4798 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAB70397.1" /db_xref="GI:2342675"	gene
CDS	/translation="MAEDTSPGQLASMTTDDIGRASRLANEIRILKESORTNLD LESVKEARKNOEKLKNKLPYLGNVILEMSPRDAEDGANLDDLSQKQCV VLKTSQTQFLPVLVGLVDPDTLKPGLNGVNDKSYLLDPLSEYDSRVKAMEVDEK PTEDNDIGLEKQIQELVEAIVLPMTHKEQFEKLGRPKPGVLLYGPFGKTLMAR ACAAQTALAGPQLQVQMFIGDGAKLVRDFAFLAKESPCIFIDEIDAIGTKRF DSEYSGREVRTMLELNQDGFSSDDRIKVIATNRADILDALMRSGRLDRKIEF PHPTREARGRLIQHRSRMNVNADVNFELARSTDDFNGAQLKAVCVCAAGMALRRDA TEVNHDFNGLIQVAKKASLANYA" complement(6633..9149) /gene="F7G19.3"	gene
CDS	complement(join(6693..6802,6892..7391,7471..7544, 7619..7721,7805..7920,8032..8127,8253..8639,8799..9149)) /gene="F7G19.3" /note="strong similarity to Oryza NADPH oxidase (gb X93301)." /codon_start=1 /evidence=not experimental /protein_id="AAB70398.1" /db_xref="GI:2342676"	gene
CDS	/translation="MLSQGLPTKDRNPVKRFAMNISYFFLENKRWIWLTIWISICI TIFTWKLYKRYKTVFVWGYCVTVAKGSAETLKFNMALILLPVCRNTITLRTSKL IGSVFPDDINPHKVAIGLAVGLHAIHLAICDFPRLHAKNVFPPMKFQDGE RPNYGMFKGTGDTGVTVMVLAVALAOSWFRNRANLPKSLKRLTGNAPWYS HLPLVIVLLVHGYFVZLSKEWYHKTWVLAVALVLLYAFERLITAFEPKAKVY LKAVYPNVLSLMSKPKYTSQGIYINCSDVSPLOWPFFSISGDDYLSVH ITLGDWTSQKLSYKVCQSPSTQSGLFIADIQANNITFRLLIDPQGAQD YRNDVLLVGLGIGATELISIRDLVANNIKNSIERGTNOHRYNAYKAYFYW TREQSLEWFSWVNAEYDSEGMEIHLNYCTSVYEEGDARSALITMLQSLHAKSG IDIVSGTRVTRHPARNRWSVFKHVAVNVHQRVGVGYCGNTCTIIGELKRLAQDFSRK TTTKFEFKENP" complement(9310..10430) /gene="F7G19.4" complement(join(9310..9492,9582..9630,9745..9904, 10061..10430)) /gene="F7G19.4" /codon_start=1 /evidence=not experimental /product="F7G19.4" /protein_id="AAB70399.1" /db_xref="GI:2342677"	gene
CDS	/translation="MESSSEGETNKISRKATGSDNPDEYVEITLEVRDETNTKVA KATLRSVLGRLKTMVKSLSFASRLDSKSGFAMFALRGLRFAIKNDVAGRWGDEVA MRQDLAVRGKLPKSKFGHCGVSESVNELPEALVRRGTTSSITKTLRFEME QITGNSFDDRLQIFPDWVDKNDLGRITGDEYKIIALSASANKLSKIKENVDEYAALI MEELDRDMLGYIRVRCLSNGYTLCTLSLKTWLNP" 11553..13929 /gene="F7G19.5" join(11553..11812,11908..12119,12289..13011,13083..13344, 13419..13929) /gene="F7G19.5" /note="Similar to Arabidopsis luminal binding protein (gb D89342)." /codon_start=1 /evidence=not experimental /protein_id="AAB70400.1" /db_xref="GI:2342678"	gene
CDS	/translation="HSSLAIEGEQKLGTVIGIDLTYSYCVGVHKNHVELIANDQG ARITPSWAFDTDLRLIEGAANQAANKPERTIFDKELIGKFDOPDQVDIRKELPY KVNKDGKPIYQVKGEEKLFSPEISAMILTKMKETAEAPLGKKIKDAVITVPAYF NDAORQATKDAIAGLVNVRINEPTGAIAIYGLDKKGSNIIYVLDGGTFDVS LTIDNGVEVLSGSDTHLGGEDFHRVMDYFIKLKKYKNDIKSDHKLKLRREC ELAKRSLSNQHVVELESFQVDFSEPLTRARPELNMDLFKTMEPVKKALDAG LKXSDIEVLVGSSTRIPKVOQMLKDFDQKPSKCTNPDEAVAYGAOVGVLSGL GGEQTKNLTLLKTDILLLOVAPLSLGIETGVGVNTNIPRNTVPIPKKSQVFTYDQ QQTTVTINVEGERSMTKNRELKGLDITLIPAPRNYFRGPIQVTFEVDANGIL QVKAEDVAKTSQSIITNDRGLTTEIEEMIREAEFEAEEDKIMKEIDARNKLET VYVNMKSTVADKELAKKISDEKQMEGVLEALEWLENVNAEKEDYDEKLKEVEL VCDPVKSYVEKTEGENEDTMMITMSYRV" complement(14446..15420) /gene="F7G19.6" complement(14446..15420) /gene="F7G19.6" /note="Similar to Glycine SRC2 (gb AB000130). ESTs gb H76869,gb T21700,gb ATT5089 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAB70401.1" /db_xref="GI:1922937"	gene
CDS	complement(14446..15420) /gene="F7G19.6" complement(14446..15420) /gene="F7G19.6" /note="Similar to Glycine SRC2 (gb AB000130). ESTs gb H76869,gb T21700,gb ATT5089 come from this gene." /codon_start=1 /evidence=not experimental /protein_id="AAB70401.1" /db_xref="GI:1922937"	gene

HKKAGAGMGLGILGAGILGGLLNGEAVSDIADMGDMGDMGDFP*
 17933..21752
 /gene="F7G19.7"
 join(17933..18538,18639..18781,18879..19029,19082..19181,
 19238..19367,19470..20115,20194..20523,20583..20618,
 20661..21032,21125..21194,21284..21369,21470..21577,
 21678..21752)
 /gene="F7G19.7"
 /note="Similar to Vicia sativa ENBP1 (gb|X95995)."
 /codon_start=1
 /evidence=not experimental
 /protein_id="AB070402.1"
 /db_xref="GI:2342679"
 /translation="MQVNEDETCDVIRNANEQTSANGIGNGESIPGIPDILRC
 RESDGQRCTAMSWADTKVCKHYIQAKRANAFKRAKQKAKRSLSIGETDIYSE
 GRQDDFELPVTSDIHYNNGLASAKSNGLEKRNKSLMRYSPETPMERSFPRVAVD

Alignment Scores: 3 48e-74 Length: 110514
 Pred. No.: 1003.50 Matches: 306
 Score: 29.30% Conservative: 65
 Percent Similarity: 24.17% Mismatches: 75
 Best Local Similarity: 42.85% Indels: 823
 Query Match: 8 Gaps: 15
 DB:

US-09-857-525C-2 (1-451) x F7G19 (1-110514)
 QY 3 GluHisLeuProMetSerProLeuGluGlyCysProThrValPheGlnAlaIleCysSer 22
 |||||:|||||
 DB 39254 GAGCATGTCGAATGTCACCGCTCGAGGGCTGCCCTACTGTGTTTCAAGTCATTTCGAC 39313
 |||||:|||||
 QY 23 LeuSerProGlyLeu 27
 |||||:|||||
 DB 39314 TTGACGCCAGGATATCATCAGGTTAATAATCTTGTACTCGACCTGTATCATGCGCAATTC 39373
 |||||:|||||
 QY 27 27
 DB 39374 TTTGTTAATTAATCCCTGCGAGAGCATTATATGTCCTTGAGTAGTATTAGCGATGCCATA 39433
 |||||:|||||
 QY 27 27
 DB 39434 ATTGTCCTTCATTAGTCTCTGGATCTTAACCTATGCAATGAGAACTCTATATACCTCT 39493
 |||||:|||||
 QY 27 27
 DB 39494 GTCAATATTGTATGACTGTGGCTGAAATTTGCAAAATTCCTTGTCTAATTGATCAAAATTT 39553
 |||||:|||||
 QY 27 27
 DB 39554 GACITTTATACATCACCATCCACATACAAAGTTGGGGCTCTCATGCTGTGTATACCTA 39613
 |||||:|||||
 QY 27 27
 DB 39614 AATCACTTCTTTTAATAAGGCTACAGTGTGTATAGATTCATATATTGCGCAGTT 39673
 |||||:|||||
 QY 27 27
 DB 39674 GGATGACTATATATACAAAGATTTAATTATTAACAAATGAGATGCTTAAGAAGCCACT 39733
 |||||:|||||
 QY 28 28
 DB 39734 AATGTTACATTATGTTATTTCACTTCATGCTCATGATAAGTTTITTTGTTGATGGGAATG 39793
 |||||:|||||
 QY 38 pArgHisAspGluArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTy 58
 |||||:|||||
 DB 39794 GCGGCACGATCAGCACCAACCATTTGTAAAGCGGAAATGGTGGAGTAGTGAATACATATTT 39853
 |||||:|||||
 QY 58 rLeuThrArgGluThrAsnGlnIleAsnThrLeuSerProSerProSerProGlySerAr 78
 |||||:|||||
 DB 39854 TATAACTGGGCCA--GATATGGTTCCTGCTGGTGTATGACCCAGACACCTGGCC--CG 39907
 |||||:|||||
 QY 78 gMetAsnMetAspValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyTh 98
 |||||:|||||

Db 39908 GTGGAATATGATGTGGAT--GATGTCTTCTTGGACAGGTAAAGAA--TCTGACCTTTC 39963
 QY 98 rValSer 100
 |||||
 Db 39964 CTTTCCAAAATGACGTGTTTATCTCGCTGCTTGTATGTTCCAGGTTTCCGAGTTGCTT 40023
 QY 100 100
 |||||
 Db 40024 CCAGTTTGGTTCTTTAATTTTGTATCCGTGTTGGTATCCTTCAGATTATTTTAACTGT 40083
 QY 100 100
 |||||
 Db 40084 TTGTCTATTAGTTCAAGTATATAATTGATTATTAGACCAATTTATCTTAATGTATTAAGCAGGA 40143
 QY 100 100
 |||||
 Db 40144 TTCTTTAAGTTCCATTAATTTGTTCTGAGACTACATTCAAATTTGAGTTTCTGACTGCTT 40203
 QY 101 101
 |||||-----GluGlyThrLeuArgValSerG1 108
 |||||:|||||
 Db 40204 TCACATGCTGATTATATTGTAGGCTGACCGCTCCAGGAAGCTGTTCTTAGGATGTCGGG 40263
 QY 108 uAlaAlaIleGlnIleSerArgCysArgValSerGluTyLeuAsnLeuHisThrCysTy 128
 |||||:|||||
 Db 40264 GGTGATTGGAGCTTCTCGTCACCGCATATCGGTTTATTTGCAACCCGCACTGCATA 40323
 QY 128 rAspLeuLeuProAspSerGlyLys 136
 |||||:|||||
 Db 40324 TGAGCTGCTCCAGATCGGCAAGTTTGGTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 40383
 QY 136 136
 |||||
 Db 40384 AATTTCGTCTTCATGCTCTATAGATTATCATGTTAAGATAAACCTTTGGTCTTCTACTCG 40443
 QY 137 137
 |||||-----ValIleAlaLeuAspIleAsnLeuProV 146
 |||||:|||||
 Db 40444 GAAATGTCPAAGGTGANTGGAATTTTACAGGTTATTCGATTCGATTCGATTCGATTCGATTCG 40503
 QY 146 allYsGlnSerPheHisIleLeuHisGlu 155
 |||||:|||||
 Db 40504 TGAAGCAAGCATTCATATATCTATGAGCAGGTTTACTTTATTTGCAATTACCTTGT 40563
 QY 155 155
 |||||
 Db 40564 GTCAGTCTCTGTTCTTCTTAGGATGTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 40623
 QY 155 155
 |||||
 Db 40624 CTAATATTGTTGCATCCAGACCCGGAAGGAAGGAAGTGAACCTTAAGTGAACAACATT 40683
 QY 156 156
 |||||-----G 156
 Db 40684 CGTCTAATATATCTTCGAGATTGTTTGGCTTTCCTGGTCTTTGCTAATTCAGATGCTTTC 40743
 QY 156 lnGlyIleProValAlaProLeuThrPaspSerPheArgGlyGlnPheValGlyLeuLeuS 176
 |||||:|||||
 Db 40744 AGGAAATCCCTTCTGCTCTCTTTGGGACCTTTGGAAGAGCCAAATTTGTTGGAGTCTCTG 40803
 QY 176 erProLeuAspPheIleLeuIleLeuArgGlu 186
 |||||:|||||
 Db 40804 GTCCACTAGACTTCATTCATATCTGAGAGAGGTGGCTTCCTNTTAAACATGTTCTTTTTC 40863
 QY 186 186
 |||||
 Db 40864 TAGACATTCCATCGAGTTACTATTGTATATTCTTAGAATTTCCAGAGGAGGAGTCTGGA 40923
 QY 186 186
 |||||
 Db 40924 TATCTCATTAATCATATAGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAA 40983
 QY 187 187
 |||||-----LeuGluThrHisGlySerAsnLeuThrG 196
 |||||:|||||
 Db 40984 TTGAGTATATAGCAAAATTTGTAATTTGACGCTTGGAACTCATCGATCAAACTTGACAG 41043


```
QY 196 luGluGlnLeuGluThrHisThrIleSerAlaTrpIysGluAlaLysArgGlnThrAsnG 216
Db 41044 AGGAGAGCTTTGAGCGCACACAAATAGACCTTGAAAGAGGGAAGGCTCATATTAGCC 41103
QY 216 lyArgAsnAspSerGlnTrpArgProGln----- 225
Db 41104 GGCAATATGATGAAGGGAGAGCCATATCCTAGGCCACTTGTTGAGTGAGATTTTCAC 41163
QY 225 ----- 225
Db 41164 AGTTGATTTGATCGCTGTACTACTAACTTTCTCTGTTGGCTGCTATCTCTGAACATAG 41223
QY 226 -----GlnHisLeuValHisAlaThr----- 232
Db 41224 CCTATCCTAGAACTTTCTATTGCGATGTTAT-ATTGAAAGAGCAGTAGTAGTATTTCTCTC 41282
QY 233 -----ProTyrGluSerLeuArgAspI 240
Db 41283 TTGCTATTAAATGTTTATATATATGCGCTATCAGGTTGGACCTATGATAATCTGAAGAGC 41342
QY 240 leAlaValIysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerS 260
Db 41343 TTGCCCTGAAATTTTGCAGAAACAAAGTGCAGCTGTTCCGTTATATATTTCTTTTGC 41402
QY 260 erAspGlySerPheProGlnLeuHisLeuAlaSerLeuSerGlyIleLeuLysCys- 279
Db 41403 AGGATGTTTCATATCCGAGTTACTGTCATCTTCTGCTATCCAGGCATATTAANAATGTA 41462
QY 279 ----- 279
Db 41463 AGGTCCCTTCATTGTGTAATAGACACTGATAAGCGGCCAATTTGGCAITTTCTCACTTT 41522
QY 279 ----- 279
Db 41523 TTAATTCATGACGGTCATATGATTTCTGCTATTTCCAAAATTTTCCCACTTTCTTCAITTTCT 41582
QY 279 ----- 279
Db 41583 CCAANTTCTGGAAGGATGGGATCTTTAGGAGGAGGGAATTAAGAACTAAAGTTGCA 41642
QY 279 ----- 279
Db 41643 ATTGACTTTACCAACTTTTACACAGTTTCCCTGTTTAAATAATTCCTTTCCGAAATTTAT 41702
QY 280 -----IleCysArgTyrP 284
Db 41703 GCCTTTATCAGACACTTAAGCTCTAAATTAAGCATAATCTTCCAGGTATATSCAGATACT 41762
QY 284 heLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuG 304
Db 41763 TTAGACATTTGCTAGCTCTTTTACTCTATCTCTTCCAGAACCCATTTGTTCAATTTCCCTGG 41822
QY 304 lySerTrpValProLysIleLysAspLeuAsnSerArgProLeuAlaMetLeuArgProA 324
Db 41823 GTAAGTGGTCCCTAGAAATCGGAAATCAAGTAGCAAAACCTTCCGCTACATTGAGACCTC 41882
QY 324 snAlaSerLeuSerSerAlaLeuAsnMetLeuValGln----- 336
Db 41883 ACGCCTCTTTGGTTCGCGCTCGCATTTATAGTTCAGAGGTAATCCCTTTTCTTATCTTTG 41942
QY 336 ----- 336
Db 41943 TTGCAGCAACATCTTATTCAAAGCTCATCTGCAACTCAGACTATCAACCTCTACATTACA 42002
QY 336 ----- 336
Db 42003 ATCAGATTTTCAGTTTAAATTAATTTTGCTAGATGGTAATATAATCAAGTAGCTGAACAC 42062
QY 336 ----- 336
Db 42063 ATGGGTATATCAAGTCAATGACAGCAACAAACCAACTTTTAGATGTTAAGCAACACTTTT 42122
QY 336 ----- 336
Db 42123 ATTATTTTATACAGATGTAATTTGCTGCTCAACATGAGTAACTAAGAGGTTTAAAAACAT 42182
QY 336 ----- 336
Db 42183 AGAAACATTTTGTCAATCTTAAAGCAGATATATTTATAGAATTTCTCGTGATCTTTTCC 42242
QY 336 ----- 336
Db 42243 CTGGTTCCTCATAGATGAAATGAATGAATGAGTCATGCTTAATTTATTTTCTCGGTGA 42302
QY 337 -----AlaGlyValSerSerIleProIleValAspAspAsnAspSerLeuLeuAspThr 355
Db 42303 TTACAGCTGAAGTCRGTTCAATTTCCGTTAGTGATGACAGACTCGCTTATTGACATAT 42362
QY 355 YrSerArg----- 357
Db 42363 ACTCTGAGAGGTAAACCTTATCCATCAATCGACATGATGCTTATGTCATTTCCATCTTTG 42422
QY 358 -----SerAspIleThrAlaLeuAlaLysA 366
Db 42423 ACGTCTTGTGTTCTGTGATGATTTTGTAAATTTGGCCAGTGATATACTGCTCTGGCTAAG 42482
QY 366 spLysValTyrThrHisValArgLeuAspGluMetThrIleHis----- 380
Db 42483 ACAAGCATACGCACAGATTTCATCTTGTATGACATGACAGTTCCACAGGTAAACCTTCTCAT 42542
QY 380 ----- 380
Db 42543 CAGAAACCAAGATCCATCATCATGCTCATCAAAATTACGAGTTTACTTAAATGTAACCTG 42602
QY 380 ----- 380
Db 42603 CCTATCATCATGAAATTTGCTGTTTCTGATTTATTTGAAACATACATCTTTTACAGAA 42662
QY 381 -----GlnA 382
Db 42663 TGAAGAAGAGAGTGAACGTTTATTAATAAGTGGAGGATAATGTGTTGTGGTTCAGG 42722
QY 382 laLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysG 402
Db 42723 CGCTACGATTTGGGCAAGATGCGAGCCGCTTATGGAATCTTCAACGGGCGAGATGTC 42782
QY 402 lnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyV 422
Db 42783 ACATGTGCTTGGCTCAGACTCTCTTGTGAAGATGATGAGCGGTTGGCGAATCCAGTA 42842
QY 422 al----- 422
Db 42843 C-ACACTATCTCTATATACAATCTTCAGCTTTTGTGTTGCTTGACAGGAACATCAGATGA 42901
QY 423 -----ArgArgValPheI 427
Db 42902 GATGTAACCTGGGTTTGGGTTTGTGTATATAAAACACACAGGTTGTAAGGAGGCTGGTA 42961
QY 427 leValGluAlaGlySerIysArgValGluGlyIleIleSerLeuSerAspIlePheLysp 447
Db 42962 TAGTGGAGCAGGAGGCAACGCTGTTGAAGGTATCAITTCATCGAGCGATGTTTTCCAAT 43021
QY 447 heLeuLeuSerLeu 451
Db 43022 TCCTGCTCGGTCTT 43035
RESULT 14
AKL121122
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023074117, full
insert sequence.
ACCESSION AKL121122
VERSION AKL121122.1 GI:37990745
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
```

ORGANISM

Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

1. The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team:
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Ohtsuki, K., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Arakawa, T., Fukuda, S.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

2. Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, K., Nagata, T., Nakahama, Y.,
Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishizaki, K.,
Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M.,
Oka, H., Osato, N., Ota, Y., Ohtsuki, K., Ohtsuki, K., Saitoh, H., Sasaki, C.,
Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,
Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y.,
Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M.,
Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K.,
Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J.,
Yokomizo, S., and Yoshimura, A.

TITLE

Collection, mapping, and annotation of 28K full-length cDNA clones
from japonica rice

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp]
Tel: 81-29-838-7007, Fax: 81-29-838-7007

This clone is one of the 32K full-length cDNA clones from japonica
rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T.,
Yamamoto, M., and Nakahama, Y.
FAIS Genome Sequencing & Analysis Group: Ohtsuki, K., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M.,
Kobayashi, M., Kodama, T., Kuroseki, T., Kusumegi, T., Lu, M.,
Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M.,
Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M.,

FEATURES

source

1. 1923
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023074117"

ORIGIN

Alignment Scores:
Pred. No.: 7,08e-59 Length: 1923
Score: 798.50 Matches: 181
Percent Similarity: 55.46% Conservative: 88
Best Local Similarity: 37.32% Mismatches: 147
Query Match: 34.09% Indels: 69
DB: 8 Gaps: 7

US-09-857-525C-2 (1-451) x AK121122 (1-1923)

Qy 10 ILeGluGly-CysPro-----ThrValPheGlnAlaIleCysSerLeuSe 24
Db 455 GTGGAGGAGGATGCCCATGGGCTCGTGGGGCCGAGTTCAGAGTGGTGTTCGATCTGCC 514
Qy 24 rProGlyIleHisGluTyrllysPhePheValAspGlyGluTrpArgHisAspGluArgG1 44
Db 515 CCCCGGGGTTATCAGTACCAGCGGTTTGGTTCAGTGGTCTGGCGCATGTGATGAGACGAA 574
Qy 44 nProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrlleThrArgGluTyraP 64
Db 575 ACCITGTGTACGGCATGAATATGGGTTGATTGACCAATGAAGTCTTGTGGATAATACACA 634
Qy 64 nGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspValAS 84
Db 635 TCCGGTT-----GTACA 646
Qy 84 pAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrVal---SerGluGlyTh 103
Db 647 GCCAAGAACTTCCATCAGAGTAGTGGATGGATGGATGGATGGATGGATGGATGGATG 706
Qy 103 rLeuArgValSer----- 107
Db 707 TATCAAGAGTTTCAATTGTGTAATTTTGTGTAATTTTATTTATGCTACCTAATAACTTAGTC 766
Qy 107 ----- 107
Db 767 TGTTCATGATTTTCCCAATTATCTCAGCAGATGCCGCCAGATCAGTTTCTCTCAAAACT 826
Qy 108 -GluAlaAla-IleGlnIleSerArgCysArgValSerGluTyrlleuAsnLeuHisThrC 127
Db 827 CAGCGTGCAGATAGCAATTTTCGCCCATCAGTCTCTGAGATA-TATTACATAATACCA 886
Qy 127 ystYrAspLeuLeuProAspSerGlyIysValIleA-aLeuAspIleAsnLeuProValL 147
Db 887 TATATGATGTTGTTCTCTGTTTCTAGCAGACAGCAGTGTTCGATGCTGCTCTCTCTGTC 946
Qy 147 yGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTrpAspSerP 167

D	b		947	AGCAAGCCCTTTAAAATCATGTCATGATGAGGGTCTTTCTTTGGTTCCTCTTTGGATGTATC	1006
Q	y		167	beArgGlyInPheValGlyLeuLeuSerProLeuAspPheIleLeuLeuLeuAraGluL	187
D	b		1007	AGCAGCAACCGTAACAGGCATGCTACCOCATCCGATTYYTGTAATTAATTTTGAGAAGC	1066
Q	y		187	eUglThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThrIleSerAlaT	207
D	b		1067	TGCAGAGAACATTCGAACCTTTGGCCATGAGGAACTTGAAATGCATTCAGTCTCTGCTT	1126
Q	y		207	rPlysGluAlaYlsArgGlnThrAsnGlyArgAsnAspSerGlnTPdrqProGlnSLH	227
D	b		1127	GGAAGAGACAAATTCGACTTTTATGCGGAGCTGATGTGTGTCGCAATTCAGAGAGGC	1186
Q	y		227	iStenValHisAlaThrProTyrgLusertLeuArgAspileAlaValYlsLeuLeuGlnA	247
D	b		1187	CNTTAATTCATGTAAAGGATCMGATAANTTAAGGATGTGCCATTTGGCTATATAATTAGAA	1246
Q	y		247	sNGlyIleSerThrValProValIleTyrserSerSerAspSgLyserPheProGlnL	267
D	b		1247	ATGMAATATCTCAGTTCCT--ATCTTTAAGCCCTCAACGGATTTCATCAGGATGCCTT	1303
Q	y		267	euleuthisLeuAlaserLeuSerGlylleuYlsCysIleCyseArsgTyrgPhelysAsnS	287
D	b		1304	TGCTTGCTCTTGCAACCTTCCAGGGATTCTGAATTTATTTGCTCAAAGCTCCAAAGAC	1363
Q	y		287	erThrGlyAsnLeuproileleuasndinProValCysSerilleProLeuGlySerTipv	307
D	b		1364	AGCTGAAGGTACTCATTTCTGCAAAACCAAGATTGTCAATNGCCATTGTGTACATGGT	1423
Q	y		307	aLPolyIllleglyAspleuasnSerArgproLeuAlaMetLeuAraProAsnAlaserL	327
D	b		1424	CACCACATCTGGCAAGCAAGCAATAGACAGCTTAGAACTTCSGACCAAGCACCTCCTC	1483
Q	y		327	eUsSerAlaLeuAsnMetLeuValGlnAlaglyWalserSerIleProIlevalAspA	347
D	b		1484	TAAATTCATGCTCGAATTGCTGTCTGAAGATAGAGTAAGCTCAATTCCTATATGTCAGC	1543
Q	y		347	sPenAspSerLeuLeuAspThryrSerArgSerAspileTheAlaLeuAlalyAspL	367
D	b		1544	ATAATGGGCTCTCCTTGATGTCTACTTCAGTGATATCATGGCTCTTAGGCNAGHTG	1603
Q	y		367	ysValTYrrHisValargLeuAspGluKethrIleHisGlnAlaleuGlnleuGlyG	387
D	b		1604	ATGCTACACTCGTATTGAGCTTGACAGGTGACGGTGGAGCATGCTCTGGAGCTGCAAT	1663
Q	y		387	lnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnAraG---CysGlnMetCysIeuA	406
D	b		1664	ACCAAGTGT-----AATGGCCGAAGACACTGTTCATACCTGCTTGA	1702
Q	y		406	rgSerAspProLeuLeuYlsValMetGluArgLeuAlaAsnProGlyValArgArgValp	426
D	b		1703	GCACTAGTACCTCTCTGAGGTTTTGGAGCAATTTGTACGCTCCAGGGTGGCGGAGTGG	1762
Q	y		426	heIlleValGluAlaGlySerlysArgValGluGlyIleIleSerLeuSerAspIlePheL	446
D	b		1763	TCGTTATTGAACCAAGGAGCAGATTGTGCAAGGANAATACTCATTTAGGGGACGCACTTA	1822
Q	y		446	ysPheLeuLeu	449
D	b		1823	CATTCTTCATT	1833

[illegible]

RESULT 15	AC146571	139089 bp	DNA	linear	HTG 12-DEC-2003
LOCUS	AC146571				
DEFINITION	Medicago truncatula clone mth2-10a2, WORKING DRAFT SEQUENCE, 4				
	unordered pieces.				
ACCESSION	AC146571				
VERSION	AC146571.6	GI:39752723			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	Medicago truncatula				
ORGANISM	Medicago truncatula				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				

Db 66608 ATTGTGGTTGTTATTTTATTGGTGTCTCAGGAGGGACATATTCATGAATAGTTTTC 66667
QY 336 ----- 336
Db 66668 TCTTGATTAGTTTGTGTGTAGTACGACAGTCTTTTGTAAATACACGGTACCTTTAACCTG 66727
QY 337 ----- -AlaGlyVal 339
Db 66728 TTTTGTGGTCTAGCTTTTGATATCCATATGCTATGTAATATGCTTCTTATTTGCAGCTAAAGTT 66787
QY 340 SerSerIleProileValAspAspAsnAspSerLeuLeuAspThrIleHisGlnAla 357
Db 66788 AGCTCAATACCAATTGTGGATGAATGATTCATTCCTTGACATATATTCACAGAGTAG 66847
QY 357 ----- 357
Db 66848 TATTTTTCCTCCCCCAATTTCCTTTTGTGTGTGANTATCATATTTATTTGAAGAT 66907
QY 358 ----- SerAspIleThrAla 362
Db 66908 CCTGAGCATATATGAATATGATGTGCTGCTATATTTTCTTTGTCAGTATATTTACTGCA 66967
QY 363 LeuAlaLysAspLysValIleThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
Db 66968 TTGGCTAAGATAAAGCATATGCTCGGATATCCCTGGATGAACAACTAATATTCATCAGGT- 67026
QY 382 ----- 382
Db 67027 ACCTACTA-TGTGATGACTTTTGAGTTTATAATAATGCAATACATTTGTTGTTTGTG 67086
QY 383 ----- -LeuGln 384
Db 67087 TTTGTTAGCAAAATTAAGTGAATCTAAATATTTGCACTACATGGTTGTAGGCATTAAAT 67146
QY 385 LeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlnArgCysGlnMetCys 404
Db 67147 TTGGACAAGATGCAAAATTCCTTATGCTTTAACAATGGGCACAGATGTCATATGCT 67206
QY 405 LeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnPro 420
Db 67207 TTGAGGTCGTGATTCAATTGCACAAAGTGATGGAGCGGTGGCTAAACCTGGTATTACTTTG 67265
QY 420 ----- 420
Db 67267 TTTACTGATATAAGCATAGCTCCCTCCACTTCTTAAATGGAAAGGTAGTTCCTT 67325
QY 421 ----- GlyValArgArgValPheIleValGluAlaGly 431
Db 67327 TATTCATCGGTTTCTGCACCTTTCAGGGGTT-AGGAGACTCGTGATTTGGAGCTGCT 67385
QY 432 SerLysArgValGluGlyIleIleSerLeuSerAspIlePheLysPheLeuLeu 449
Db 67386 ACCAAGCGTGTGGAGGATATTTCAITTAAGTATGTTTTCAGATTCTTATTG 67439

Search completed: July 9, 2004, 10:15:08
Job time : 5816 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2004, 05:05:39 ; Search time 609 seconds

(without alignments)
3146.038 Million cell updates/sec

Title: US-09-857-525C-2

Perfect score: 2342

Sequence: 1 TREHLPMSIECPVFEQAI.....SKRVEGIISLSIDIEKFLLSL 451

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q-/cgn2_1/USPTO.spool/US09857525/runat_07072004_161017_283/app_query.fasta_1.647
-DB=N_Geneseq_29Jan04 -QPMF=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human4.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09857525_QCGN_1_1_470_@runat_07072004_161017_283 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*

1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	1576	3 Aaa51347	Aaa51347 Z. mays s
2	2085	89.0	2149	3 Aaa51348	Aaa51348 Rice suc
3	2031	86.7	1632	3 Aaa51353	Aaa51353 Wheat suc
4	1581	67.5	2538	3 Aaa51351	Aaa51351 Soybean s
5	1525	65.1	2163	3 Aaa51350	Aaa51350 Soybean s
6	923	39.4	538	3 Aaa51354	Aaa51354 Maize suc
7	530	22.6	514	3 Aaa51357	Aaa51357 Wheat suc
8	460	19.6	282	6 ABL70761	ABL70761 Corn tass

9	392	16.7	702	3 Aaa51349	Aaa51349 Rice suc
10	373	15.9	1167	6 AAL48614	Aal48614 Human ins
11	373	15.9	1435	2 AAX06882	Aax06882 Disease a
12	373	15.9	2223	4 AAHL4839	Aahl4839 Human cdx
13	373	15.9	2223	6 AAL48615	Aal48615 Human ins
14	367	15.7	1578	6 AAL48616	Aal48616 Human ins
15	367	15.7	1578	6 ABK84324	Abk84324 Hmar: cdx
16	367	15.7	1578	7 ACA56515	Aca56515 Human sig
17	367	15.7	1691	7 AAC98774	Aac98774 Human pan
18	364	15.5	1576	2 AAT85927	Aat85927 Mammalian
19	361.5	15.4	1905	6 ABS51409	Ab551409 cDNA enco
20	361.5	15.4	3261	4 ABL18857	Ab118857 Drosophi
21	340	14.5	1647	5 AAH43685	Aah43685 PRKAG3 CD
22	340	14.5	2109	4 AAD03296	Aad03296 Human AMP
23	340	14.5	2115	4 AAD03320	Aad03320 Human AMP
24	340	14.5	2314	6 ABQ61173	Abq61173 AMP kinas
25	340	14.5	2781	7 AAD49456	Aad49456 Human kin
26	336	14.3	1774	6 ABL39755	Ab139755 Human NS
27	335.5	14.3	1873	6 AAD36460	Aad36460 Pig PRKAG
28	335.5	14.3	1873	6 AAD36459	Aad36459 Pig PRKAG
29	334.5	14.3	1857	4 AAD03295	Aad03295 Pig AMPK
30	334.5	14.3	1873	4 AAD03319	Aad03319 Pig AMPK
31	334.5	14.3	1873	6 AAD36458	Aad36458 Pig PRKAG
32	334.5	14.3	1873	6 AAD36457	Aad36457 Pig PRKAG
33	334.5	14.3	1873	6 AAD36456	Aad36456 Pig wild-
34	334.5	14.3	2022	4 AAD03321	Aad03321 Sus scrof
35	309	13.2	39651	4 ABL18856	Ab118856 Drosophi
36	303.5	13.0	969	7 ACC60984	Acc60984 Gene sequ
37	302	12.9	1467	5 AAS84265	Aas84265 DNA enco
38	302	12.9	1467	9 ADE07539	Ade07539 Novel cod
39	302	12.9	1467	9 ADE09727	Ade09727 Novel DNA
40	269.5	11.5	1266	3 Aaa51352	Aaa51352 Soybean s
41	226.5	9.7	542	3 Aaa51355	Aaa51355 Rice suc
42	223.5	9.5	178	6 ABL75404	Ab175404 Corn tass
43	211.5	9.0	735	4 AAB07561	Aab07561 Human cdx
44	184.5	7.9	2303	5 AAS84267	Aas84267 DNA enco
45	183	7.8	765	6 ABL89686	Ab189686 Human pol

ALIGNMENTS

RESULT 1

AAA51347

ID AAA51347 standard; cDNA; 1576 BP.

XX

AC AAA51347;

XX

DT 26-SEP-2000 (first entry)

XX

DE Z. mays sucrose non-fermenting 4 protein cDNA from clone cskic.pk001.cl5.

XX

KW Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;

KW carbon catabolite repression; development; nitrogen partitioning; ss.

XX

OS Zea mays.

XX

PH Key Location/Qualifiers

CDS 3..1358

FT /*tag= a

FT /label= SNF4

FT /product= "Sucrose_non-fermenting_4_protein"

FT /partial

XX

DN W0200036116-A2.

XX

PD 22-JUN-2000.

XX

PF 15-DEC-1999; 99WO-US029825.

XX

PR 16-DEC-1998; 38US-0112564P.

XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX

PI Allen SM, Heppard EP, Miao G, Weng Z;
 XX WPI: 2000-431594/37.
 DR P-PSDB; AAY96783.
 XX
 ET New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 XX and development.
 XX
 PS Claim 2; Page 31; 48pp; English.
 XX

CC This cDNA encodes maize (Zea mays) sucrose non-fermenting 4 (SNF4)
 CC protein which is involved in carbon catabolite repression in plants and
 CC seeds. The cDNA was isolated, based on similarity to SNF4 proteins from
 CC Arabidopsis thaliana and Saccharomyces cerevisiae, from clone
 CC cskic.pk31.c15 prepared from corn unpollinated developing silk 24 hours
 CC after emergence. The polynucleotides are used in plants to control carbon
 CC and nitrogen partitioning pathways during plant growth and development.
 CC The catabolite repression proteins would facilitate studies for better
 CC understanding the mechanism of catabolite repression in plants and could
 CC provide genetic tools to enhance or otherwise alter the accumulation of
 CC carbohydrates, lipids and proteins during plant growth and development
 XX
 SQ Sequence 1576 BP; 418 A; 336 C; 362 G; 460 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,45e-223 Length: 1576
 Score: 2342.00 Matches: 451
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DE: 3 Gaps: 0

US-09-857-525C-2 (1-451) x AAA51347 (1-1576)

Qy 1 ThrArgGluHisLeuProMetSerProLeuGluGlyCysProThrValPheGlnAlaIle 20
 Db 3 ACGAGGGAACATTTGGCCGATGTCCTATCGAAGGCTGCCACCTGTATTTTCAGGCTATT 62
 Qy 21 CysSerLeuSerProGlyIleHisGluCysPheValAspGlyGluTrpArgHis 40
 Db 63 TGCAGTCTGCTCTCTGGGATTACAGGACAAATTCCTTTGTGGAGCGGGAATGGCGGCAT 122
 Qy 41 AspGluArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTrpLeuThr 60
 Db 123 GATGAGCGTCAACCTACCATATCTGGGAGTTTGGCATAGTTAAACACTTACTTGACA 182
 Qy 61 ArgGluTrpAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsn 80
 Db 183 AGGGAATATAACCAATAAACCCTTATCAAGTCCAGCACACCTGGGAAGCAGGATGAAC 242
 Qy 81 MetAspValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSer 100
 Db 243 ATGGATGTGGATATGAAATTTTCAAGTACGTTAGTTGTAGTGGCACCCTTTTCA 302
 Qy 101 GluGlyThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGlu 120
 Db 303 GAAGGTACTCTGAGAGTTTCAGAGGCTCAATAACAATATCTAGGTGCGGTGTTCTGAA 362
 Qy 121 TyrLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyIleValIleAlaLeu 140
 Db 363 TATCTGAATTTGCATACATGCTATGATTTACTCCAGATTTCTGGCAAGGTTATTGCCCTA 422
 Qy 141 AspIleAsnLeuProValIleGlnSerPheHisIleLeuHisGluGlnGlyIleProVal 160
 Db 423 GACATTAATTTACTGTGAGCATCATCTCCATATTCTCCATGACAGGGGATTTCTGTGA 482
 Qy 161 AlaProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPhe 180
 Db 483 GCTCCTCTCTGGGACTCATTCAGAGGTCAATTTGTTGCTCTCTTAGCCCATTTGGATTTC 542
 Qy 181 IleLeuIleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGlu 200

Db 543 ATACTCATATTGGGGAGCTAGAACTCATGGCTCGAACTTGACAGAAGAGCAGCTTGAA 602
 Qy 201 ThrHisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSer 220
 Db 603 ACACACACTATATCTGATGGAAGAGGCTAAGCGGCAAACTAATGGAAGAATGATAGT 662
 Qy 221 GlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGlnSerLeuArgAspIle 240
 Db 663 CAGTGGCGACCGCAACAGCATCTAGTGCATGCCCTTATGAGTCTCTTGAGGGAATT 722
 Qy 241 AlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTrpSerSerSer 260
 Db 723 GCAGTAAGCTTTTCGAAATGGCATTCTACAGTCCGAGTTATTTTTCATCATCATCA 782
 Qy 261 AspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIle 280
 Db 783 GATGATCATTTCCCGCAGTTATTGCATCTTGCATCATTCTGGAATTTTGAATGTATT 842
 Qy 281 CysArgGlyPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSer 300
 Db 843 TGTAGATACTTCAAAACACTCAACTGGTAATTTGGCTTATTTCTGAACCAACAGTGTGCTCC 902
 Qy 301 IleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMet 320
 Db 903 ATTCCGCTGGGTTCCTGGGTCCGAAATTTGGTGATCTGAAACAGTCTGCTCATTTGGCTATG 962
 Qy 321 LeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSer 340
 Db 963 TTGGACCTTAATGCTCTAGCTCTGCTTAACTGTTGGTTCAGCTGGAGTAAGC 1022
 Qy 341 SerIleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIle 360
 Db 1023 TCATPACCAATTTGTGATGACACGACTCCCTGCTTGACACTTACTCTAGAAGTGACATC 1082
 Qy 361 ThrAlaLeuAlaLysAspLysValTrpThrHisValArgLeuAspGluMetThrIleHis 380
 Db 1083 ACAGCCCTAGCAAAAGACAAGGCTACACATGTTGCGTGGATGAGATGACCATTCAC 1142
 Qy 381 GlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArg 400
 Db 1143 CAGGCTTTTGCAGCTTGGACAAGATGCCAATACACCTTTTGGATTCTTTAAACGCCAGAGA 1202
 Qy 401 CysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnPro 420
 Db 1203 TGCCAGATGTGCTCCGCTCGATCTGATCTTGTCTGAAGGTGATGGAGCGACTGGCTAATCT 1262
 Qy 421 GlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSer 440
 Db 1263 GGGGTGCGCGGGTGTTCATTGTGGAAGCTGGAGCAAAACGTTGGAGGCAATCATATCA 1322
 Qy 441 LeuSerAspIlePheLysPheLeuLeuSerLeu 451
 Db 1323 CTAAAGTGAATTTTCAAGTTCTTGTCTGAGCTTG 1355

RESULT 2

AAA51348

ID AAA51348 standard; cDNA; 2149 BP.

XX AAA51348;

AC AAA51348;

XX 26-SEP-2000 (first entry)

DE Rice sucrose non-fermenting 4 protein cDNA from clone r10n.pk083.i18.

KW Rice; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 XX carbon catabolite repression; development; nitrogen partitioning; ss.

XX Oryza sativa.

XX Key Location/Qualifiers
 XX CDS 185..1666
 FT /*tag= a

```
/product= "Sucrose non-fermenting 4 protein:"
```

WO200036116-A2.

22-JUN-2000.

15-DEC-1999; 99WO-US029825.

16-DEC-1998; 98US-0112564P.

(DUPO) DU PONT DE NEMOURS & CO E I.

Allen SM, Heppard EP, Wiao G, Weng Z;

WPI; 2000-431594/37.

2-PSDB; AA96784.

New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins involved in carbon catabolite repression in plants and seeds, useful for controlling carbon and nitrogen partitioning pathways during plant growth and development.

Claim 2: Page 33-34; 48pp; English.

This cDNA encodes rice (*Oryza sativa*) sucrose non-fermenting 4 (SNF4) protein which is involved in carbon catabolite repression in plants and seeds. The cDNA was isolated, based on similarity to SNF4 proteins from *Arabidopsis thaliana* and *Saccharomyces cerevisiae*, from clone r10n.pK083.118 prepared from rice 15 day old leaf. The polynucleotides are used in plants to control carbon and nitrogen partitioning pathways during plant growth and development. The catabolite repression proteins would facilitate studies for better understanding the mechanism of catabolite repression in plants and could provide genetic tools to enhance or otherwise alter the accumulation of carbohydrates, lipids and proteins during plant growth and development.

Sequence 2149 BP: 523 A; 492 C; 501 G; 633 T; 0 U; 0 Other;

Alignment Scores:

Argument Scores:		
Pred. No.:	1,66e-197	Length: 2143
Score:	2085.00	Matches: 391
Percent Similarity:	94.8%	Conservative: 34
Best Local Similarity:	87.2%	Mismatches: 23
Query Match:	89.03%	Indels: 0
Na.	3	Gaps: 0

US-09-857-525C-2 (1-451) x AAA51348 (1-2149)

Qy	3	GlutHisLeuProMetSerProIleGluGlyCysProThrValPheSlnAlaIleCysSer	22
Db	320	GAACATTTGCCAAATGTCCTCTGTTGAAGGTTGTCCACCGGTTTTTTCAGGCCCATATGCAGT	379
Qy	23	LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu	42
Db	380	TTGTCCCTTGGATTCATCAGTACAAAGTTTGTGTGGATGGGGAATGGCGCACGATGAG	439
Qy	43	ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu	62
Db	440	CGGCAACTTACCATTAACAGGAGACTATGGTGTGGTTAACACATTATGCTTGATAGGGAC	499
Qy	63	TyrAsnGlnIleAsnThrLeuSerProSerThrProGlySerArgMetAsnMetAsp	82
Db	500	TTTGACCAGATAAATACAAATATTGAGCCCTAGTACACCTGGAGTAAGGATGAACATGGAT	559
Qy	83	ValAspAsnGluAsnPheClnArgThrValThrLeuSerAspGlyThrValSerGluGly	102
Db	560	GTGGCAACACCAATTTCAACGGTACTGTTCTTGTCTGATGGCAATTATCAGGAAGGT	619
Qy	103	ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu	122
Db	620	CCTCAGAGAAATTCAGAGGCAGCTATACAGATCTCTAGGTGTCTGTAGCGGATTTCTTG	679
Qy	123	AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle	142

DT 26-SEP-2000 (first entry)
 XX Wheat sucrose non-fermenting 4 protein cDNA from clone wll.pk0002.b3.
 DE Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;
 KW carbon catabolite repression; development; nitrogen partitioning; ss.
 XX Triticum aestivum.
 OS
 XX Key Location/Qualifiers
 FH 13.1399
 FT /*tag= a
 FT /product= "Sucrose_non-fermenting_4_protein"
 XX
 XX WO200036116-A2.
 XX 22-JUN-2000.
 XX 15-DEC-1999; 99WO-US029825.
 XX 16-DEC-1998; 98JS-0112564P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Allen SM, Heppard BP, Miao G, Wang Z;
 XX WPI; 2000-431594/37.
 DR P-PSDB; AAY96789.
 XX
 XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins
 PT involved in carbon catabolite repression in plants and seeds, useful for
 PT controlling carbon and nitrogen partitioning pathways during plant growth
 PT and development.
 XX
 PS Claim 18; Page 43; 48pp; English.
 XX
 CC This cDNA encodes wheat sucrose non-fermenting 4 (SNF4) protein which is
 CC involved in carbon catabolite repression in plants and seeds. The cDNA
 CC was isolated, based on similarity to SNF4 proteins from Arabidopsis
 CC thaliana and Saccharomyces cerevisiae, from library wll prepared from
 CC wheat leaf from 7 day old seedling. The polynucleotides are used in
 CC plants to control carbon and nitrogen partitioning pathways during plant
 CC growth and development. The catabolite repression proteins would
 CC facilitate studies for better understanding the mechanism of catabolite
 CC repression in plants and could provide genetic tools to enhance or
 CC otherwise alter the accumulation of carbohydrates, lipids and proteins
 CC during plant growth and development
 XX
 SQ Sequence 1632 BP; 431 A; 370 C; 389 G; 442 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,69e-192 Length: 1632
 Score: 2031.00 Matches: 386
 Percent Similarity: 93.74% Conservative: 33
 Best Local Similarity: 86.35% Mismatches: 26
 Query Match: 86.72% Indels: 2
 DB: 3 Gaps: 1
 US-09-857-525C-2 (1-451) x AAAS1353 (1-1632)
 QY 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerIleu 23
 DB 2 CACGAGCGATGTCTCCAGTCGAAGGTTCGCCACTGTATTCAAGCTATCTGCAACTTA 61
 QY 24 SerProGlyIleHisGluTyrLysPheValAspGlyGluTrpArgHisGluArg 43
 DB 62 CCTCCAGGGATTTATCAGTACAGTTCAATGTGACGGCGAGTGAGGCGATGACGAGGG 121
 QY 44 GlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGluTyr 63
 DB 122 CAACCTACTATTAACCTGGAGAGATGGGGTGTAAACACTTTATCTTGACAGGGAATTT 181
 QY 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83

Db 182 GACCACATAAATCTGTACTGAGCCCACTACACTGGAGGAGG-----ATGATGTG 235
 QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
 Db 236 GACAGTGACAGTTTTCACAGGAATGGTTCGTTCGGATGGTGCCTTCAGGAGGTTCT 295
 QY 104 LeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsn 123
 Db 296 CCAAGATCTCAGAGCGTCTATACAGATCTCTAGGTGTGCTGTGCTGAGTATCTGAAT 355
 QY 124 LeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
 Db 356 GCGCATACAGGCTATGACCTACTACCAATCTCTGAAAGGTCAATGCTCTGCACATTAA 415
 QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
 Db 416 TTACCTGTGAAGCAATCTTTCATATCTCCATGAACAGGGGATTCCTGTGCTCTCTG 475
 QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183
 Db 476 TGGGATTCATTCAGGGGTCAAGTTGTGGCTCTGAGCCCACTGGATTTTATCTTATA 535
 QY 184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnGlnLeuGluThrHisThr 203
 Db 536 TTGAGAGAGCTGGAAACTCATGGCTCAAACTCTGACAGAGGACACGCTTGAACACACACT 595
 QY 204 IleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArg 223
 Db 596 ATATCTGCTGCAAGAGGCTAAGCGCAAACTTATGGAAGAAATGATGGACACTTACA 655
 QY 224 ProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLys 243
 Db 656 TCRAATCAGCATCTAGTCATGCCACCCCTTATGAATCTTGAGGGGTATTGCCATGAAA 715
 QY 244 LeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGlySer 263
 Db 716 ATACTGGAACCTGGCATTTCTACAGTCCCAATCATCTATTCTCATCATCAGATGGATCG 775
 QY 264 PheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyr 283
 Db 776 TTTCGCGAGCTGTGTCATCTTCATCCCTTTTCAGGAATTTTGAAGAATGATCTGTAGATAC 835
 QY 284 PheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeu 303
 Db 836 TTCAGAACTCCACTGGTAGTTGCGGATTTCTAAACCAACCAAGATGATGCTCAATCCGCTG 895
 QY 304 GlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgPro 323
 Db 896 GGTACTCTGGGTTCCAAAAATTTGGTGAACCAAAATGGTTCATCCATTGGCTATGTTGGCGCT 955
 QY 324 AsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIlePro 343
 Db 956 AATACATCTCTAGCTCTGCGCTTAACCTTATGTTGTTCAGCTGAGGTAGTTCATATCCC 1015
 QY 344 IleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeu 363
 Db 1016 ATTGTGATGATACGACTCGCTGATCGACACACTCCAGAGGTGACATCACAGCTCTA 1075
 QY 364 AlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeu 383
 Db 1076 GCGAAGAACAGAGGTCTACACCCATATCGCCCTAGATGAGATGACCATTCATCAGGCCCTG 1135
 QY 384 GlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMet 403
 Db 1136 CAGCTCGGGCAAGACGGCAATTCACCTTTGGACCTTTTCAATGGTCAAGATGCCAGATG 1195
 QY 404 CysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyValArg 423
 Db 1196 TGTCTCAGCTGACCCCTTTCGAGGTTATGAGAGATTTGGCTAATCTCTGGGGTGGCT 1255
 QY 424 ArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerSerAsp 443

1256 CCGGTGTCATCGTGAGGCTGGCAGCAGCGAGTGGAAAGCGTATATCATGCTGAGCGAC 1315

DE	1256	CGCGTTCATCGTGAGGCTGGCAGCAGCGAGTGGAAAGCGTATATCATGCTGAGCGAC	1315
QY	444	IlePheLysPheLeuLeuSer 450	
DB	1316	ATATTCAAGTGTCTGGAGC 1336	
RESULT 4			
AAA51351			
ID	AAA51351	s-standard; cDNA; 2538 BP.	
XX	AA51351;		
AC	AA51351;		
XX	26-SEP-2000	(first entry)	
DE	Soybean	sucrose non-fermenting 4 protein cDNA from clone sf11.pk0004.b4.	
KW	Soybean;	sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;	
KW	carbon	catabolite repression; development; nitrogen partitioning; ss.	
XX	Glycine	max.	
XX	Key	Location/Qualifiers	
FF	CDS	395..1873	
FF		/*tag= a	
FF		/product= "Sucrose non-fermenting 4 protein"	
XX	W020036116-A2.		
XX	22-JUN-2000.		
XX	15-DEC-1999;	99WO-US029825.	
XX	16-DEC-1998;	98US-0112564P.	
XX	(DUPO)	DU FONT DE NEMOURS & CO E I.	
XX	Allen SM,	Heppard EP, Miao G, Weng Z;	
XX	WPI;	2000-431594/37.	
XX	P-PSDB;	AA936787.	
XX	New	nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins	
XX	involved	in carbon catabolite repression in plants and seeds, useful for	
XX	controlling	carbon and nitrogen partitioning pathways during plant growth	
XX	and	development.	
XX	Claim 2;	Page 39-40; 48pp; English.	
XX	This	cDNA encodes soybean (Glycine max) sucrose non-fermenting 4 (SNF4)	
XX	protein	which is involved in carbon catabolite repression in plants and	
XX	seeds.	The cDNA was isolated, based on similarity to SNF4 proteins from	
XX	Arabidopsis	thaliana and Saccharomyces cerevisiae, from library sf11	
XX	plants	to control carbon and nitrogen partitioning pathways during plant	
XX	growth	and development. The catabolite repression proteins would	
XX	facilitate	studies for better understanding the mechanism of catabolite	
XX	repression	in plants and could provide genetic tools to enhance or	
XX	otherwise	alter the accumulation of carbohydrates, lipids and proteins	
XX	during	plant growth and development	
XX	Sequence	2538 BP; 670 A; 463 C; 571 G; 834 T; 0 U; 0 Other;	
Alignment Scores:			
Pred. No.:	3.94e-147	Length: 2538	
Score:	1581.00	Matches: 305	
Percent Similarity:	81.15%	Conservative: 61	
Best Local Similarity:	67.63%	Mismatches: 75	
Query Match:	67.51%	Indels: 10	
DB:	3	Gaps: 4	
US-09-857-525C-2 (1-451) x AAA51351 (1-2538)			
QY	3	GlutHisLeuProMetSerProiledGlyCysProThrValPheGlnAlaIleCysSer 22	

DB	530	GAGCTTCTACCGATGTCGCGAGGTTGTCACACTGTGTTCAAGTATTATAAC	589
QY	23	LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu	42
DB	590	TTGCCACCCGGTTACCATCAGTACAAAGTTTTTTTGTGATGGAGATGCGCGCATGAA	649
QY	43	ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuArgGlu	62
DB	650	CATCAACCTTATGTACCTACCGAATATGGATAGTCAACACTGTCTTATTGGCCACTGAT	709
QY	63	TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp	82
DB	710	CCTAACTACATGCTGTTTAA-----CCTCCAGACGTTGCTTCTGGAAATAGCATGGAT	763
QY	83	ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly	102
DB	764	GTGGATTAATGATGCTTTCCCGAATGCGCGGTGGACGATGGTACTTTGAGTGAGGTG	823
QY	103	ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrIleu	122
DB	824	CTGCCAAGATATCAGATACCTGATGTACAAATATCCCGTCAGCGTATTCTTGCATTTCTA	883
QY	123	AsnLeuHisThrCysTyrAspLeuProAspSerGlyLysValIleAlaLeuAspIle	142
DB	884	TCTTACACACCGCTTATGAATTACTTCCCGAGTCAGGCAAGGTGTGTCTTGGATGTT	943
QY	143	AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro	162
DB	944	GATTCACAGTGAACAGGCAATTCATATATTGATGAGCAGGAGTTTTCATGGCTCCT	1003
QY	163	LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu	182
DB	1064	CTTTGGGACTTCTGCAAGGGGCAATTTGTTGGTCTTCTTAGTCTTCCGATTTATTATTA	1063
QY	183	IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHis	202
DB	1064	ATTTTAAGAGAGCTGGGGAAATCATGGATCCCAATCTGACGAGAGGAGCTTGAACACAT	1123
QY	203	ThrIleSerAlaTrpLysGluAlaLys-----ArgGlnThrAsnGlyArgAsn	218
DB	1124	ACCATATCAGCTTGGMAAGAGAAATCGTATCTAAATAGACAGCAATGACATGCA	1183
QY	219	AspSerGlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArg	238
DB	1184	ACTGCATTTTCAAGA-----TGTTTATCATGTCAGGCGCCATATGATAATCTGAAA	1234
QY	239	AspIleAlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSer	258
DB	1235	GATATTGCGATGAAGATCTTGCMAAGAGAGGTTTCACTGTTCTTATTATCCATTCATCT	1294
QY	259	SerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLys	278
DB	1295	TCTGAGATGCTTCATTTCCACAGTTACTACATCTTGCTTCACTTCCAGGATACTTAAA	1354
QY	279	CysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProVal	298
DB	1355	TGCATTTGTAGGTATTTTAGGCACCTGCTCTAGTTCCCTTGCCCTGTACTTCAACTTCAATC	1414
QY	299	CysSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeu	318
DB	1415	TGTGCATACCTGTGGCGACGTGGGTGGCCAAAATTTGGGAATCAAAATCGCGGCTCTTA	1474
QY	319	AlaMetLeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGly	338
DB	1475	GCNATGTTGAGACCAACCGCTTCTTTCGCGTCAGCCCTAAATTTATTAGTTCAAGCCCAA	1534
QY	339	ValSerSerIleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSer	358
DB	1535	GTAAAGCTCATATACCAATATTTGATGATAGTACTATTATTGGATATATATCTGTCGGAGT	1594
QY	359	AspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThr	378

Db	1595	GACATACAGCTTTGGCAAGAAAGAGAGCATATACATATTAATCTTGACGAATGACT	1654
Qy	379	IleHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGly	398
Db	1655	GTTTCATCAGGCTTTGAGCTTGGCCAGGACCATATAGTCCCTATGAG---CTTAGAAGT	1711
Qy	399	GlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAla	418
Db	1712	CAAAAGATGTCAGATGTTGGTGTCTGATCTCTGCATAAAGTGATGGAACGCTTGSCA	1771
Qy	419	AsnProGlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIle	438
Db	1772	AATCCAGGTGTCAGCGGCTTTGTCACTGTGAAGCTGGGACGCGGTAGAGGCATT	1831
Qy	439	IleSerLeuSerAspIlePheLysPheLeuLeu 449	
Db	1832	GTTTCATTCAGTGACATATTCAGATTCCTTCATT 1864	
RESULT 5			
ID	AAA51350		
XX	AAA51350	standard; cDNA; 2160 BP.	
AC	AAA51350;		
DT	26-SEP-2000	(first entry)	
DE	Soybean	sucrose non-fermenting 4 protein cDNA from clone ses4d.pk040.bl.	
KW	Soybean;	sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;	
KW	carbon	catabolite repression; development; nitrogen partitioning; ss.	
OS	Glycine	max.	
PH	Key	Location/Qualifiers	
FT	185..1633		
FT	/*tag= a		
FT	/product= "sucrose_non-fermenting_4_protein"		
XX	WO200036116-A2.		
PN	22-JUN-2000.		
XX	15-DEC-1999;	99WO-US029825.	
XX	16-DEC-1998;	98US-0112564P.	
XX	(DUPO)	DU PONT DE NEMOURS & CO E I.	
PI	Allen SM,	Heppard EP, Miao G, Weng Z;	
XX	WPI; 2000-431594/37.		
DR	P-PSDB; AAY96786.		
XX	New	nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins	
PT	involved	in carbon catabolite repression in plants and seeds, useful for	
PT	controlling	carbon and nitrogen partitioning pathways during plant growth	
PT	and	development.	
XX	Claim 2;	Page 36-37; 48pp; English.	
PS	This	cDNA encodes soybean (Glycine max) sucrose non-fermenting 4 (SNF4)	
CC	protein	which is involved in carbon catabolite repression in plants and	
CC	seeds.	The cDNA was isolated, based on similarity to SNF4 proteins from	
CC	Arabidopsis	thaliana and Saccharomyces cerevisiae, from library ses4d	
CC	prepared	from soybean embryonic suspension 4 days after subculture. The	
CC	polynucleotides	are used in plants to control carbon and nitrogen	
CC	partitioning	pathways during plant growth and development. The catabolite	
CC	repression	proteins would facilitate studies for better understanding the	
CC	mechanism	of catabolite repression in plants and could provide genetic	
CC	tools	to enhance or otherwise alter the accumulation of carbohydrates,	
CC	lipids	and proteins during plant growth and development	
XX	Sequence	2160 BP; 588 A; 392 C; 485 G; 693 T; 0 U; 2 Other;	

Alignment Scores:	1.06e-141	Length:	2160
Pred. No.:	1525.50	Matches:	303
Score:	80.18%	Conservative:	57
Percent Similarity:	67.48%	Mismatches:	72
Best Local Similarity:	65.14%	Indels:	17
Query Match:	3	Gaps:	6
DB:			
US-09-857-525C-2 (1-451) x AAA51350 (1-2160)			
Qy	3	GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer	22
Db	317	GAACCTCTACAAATGTCGCTGTGAAGCTGCCCAACTGTGTTTCACTTATTATCATAGC	376
Qy	23	LeuSerProGlyIleHisGluThrLysPhePheValAspGlyGluTrpArgHisAspGlu	42
Db	377	TTGGTACCTGGTCATCATCAGTACCAAGTTTTTGTGTGAGAGATGGCGGATGATGAC	436
Qy	43	ArgGlnProThrIleSerGlyGluPheGlyIleValAlaAsnThrLeuTyrLeuThrArgGlu	62
Db	437	CTTCAACCTTGTGAATCTGGAGATATGGAATGTTTAACTGTTTCACTTGGCTACTGAT	496
Qy	63	TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsp	82
Db	497	CCTAATATTTTACCTGTTTAACTCCAGACATAGTCTCTGGATCT-----AACATGGAT	550
Qy	83	ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly	102
Db	551	GTCGACACACGAGCTTTTCGACGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG	610
Qy	103	ThrLeu---ArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyr	121
Db	611	TTATTGCCAAGATATCTGATGTTGATATACAGACCTCTGTCAGCGTATTTCTGCAATC	670
Qy	122	LeuAsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAsp	141
Db	671	CTATCTATGAGTACAGCATATGAAATTACTCTCCGTGATCAGGCAAGGTTGTACTTGGAT	730
Qy	142	IleAsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlnIleProValAla	161
Db	731	GTTGATCTGCTGTGAACAGGCAATTCATATCTGATGATGATGATGATGATGATGATGAT	790
Qy	162	ProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIle	181
Db	791	CCTCTCTGGGACATCTGCAAGGCGCAGTTTGTGGAGTCTTCTAGTCCCTTGGATTTATT	850
Qy	182	LeuIleLeuArgGluLeuGlnThrHisGlySerAsnLeuThrGluGluGlnLeuGluThr	201
Db	851	TTAATTTTAAAGAGCTCGGAAATCATGGGTCCAACTCAACAGAGAGAGGAGCTTGAACA	910
Qy	202	HisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln	221
Db	911	CATACCATATCAGCTTGGAAAGGGAAGA-----940	
Qy	222	TrpArgPro---GlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIle	240
Db	941	TGGACAGGATTACAAATGTTTATCCGTGCGAGGCGCCATACAGTAATTTTGAAGAGATT	1000
Qy	241	AlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSer	260
Db	1001	GCTGTGAAGATCCCTGCAACATGAAATTTCAACAGTCTCTATTATTCAT-----TCAGAA	1054
Qy	261	AspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIle	280
Db	1055	GATGTTTCAATTCACAGCTACTACATCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1114
Qy	281	CysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSer	300
Db	1115	TGCAGTATTTTAGGAATTTGTTAGTTCATGCTTCTTCTTCTTCTTCTTCTTCTTCTT	1174
Qy	301	IleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMet	320

Db 1175 ATCCCTGTGGCAGCTGGTCCCAAAATGGGAGTCAAAATGCCGGCTCTAGCAATG 1234

QY 321 LeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSer 340

Db 1235 TTGAGACCAATGCTTCACTTACTTTCAGCCCTAAACTTATTAGTCAAGCTCAAGTAAGT 1294

QY 341 SerTleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIle 360

Db 1295 TCAATACCAATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1354

QY 361 ThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHis 380

Db 1355 ACAGCTTTGCCAAGGACAGAACTTATACCAATATTAATCTTGATGAATGATGTTTCAT 1414

QY 381 GlnAlaLeuGlnLeuGlyGlnAlaAsnThrProPheGlyPhePheAsnGlyGlnArg 400

Db 1415 CAGGCATTGCATTTGGGCCAGGATCTTATTAATCTTATGAG---CTGAGTTGTCAAGA 1471

QY 401 CysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnPro 420

Db 1472 TGTGAGATGTTTGGCACTGATCTCTGCATTAAGTGAACGTTTGGCAAGTCCA 1531

QY 421 GlyValArgArgValPheIleValGlnAlaGlySerLysArgValGlnGlyIleIleSer 440

Db 1532 GGTGTGAGCGGCTGTGAATTTGGAGCTGGCAGCAAGCGGTAGAAGGCATCATGCA 1591

QY 441 LeuSerAspIlePheLysPheLeuLeu 449

Db 1592 CTGAGTGACATATCAACTTCTCCTT 1618

RESULT 6

AAAS1354

ID AAAS1354 standard; cDNA; 538 BP.

XX AC AAAS1354;

XX AC AAAS1354;

DT 26-SEP-2000 (first entry)

XX Maize sucrose non-fermenting 4 protein cDNA from clone csk1c.pk001.c15.

XX Maize; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;

XX carbon catabolite repression; development; nitrogen partitioning; ss.

XX Zea mays.

Key Location/Qualifiers

CDS 347..523

FT /*tag= a

FT /product= "partial_SNF4_protein"

FT 494..496

FT /*tag= b

FT /note= "Encodes Xaa which is not defined"

XX WO200036116-A2.

XX 22-JUN-2000.

XX 15-DEC-1993; 99WO-US029825.

XX 16-DEC-1993; 98US-0112564P.

XX (DUPO) DU PONT DE NEMOURS & CO B I.

XX Allen SM, Heppard EP, Miao G, Weng Z;

XX WPI; 2000-431594/37.

XX P-PSDB; AAY96789.

XX New nucleic acids encoding sucrose non-fermenting 4 (SNF4) proteins

XX involved in carbon catabolite repression in plants and seeds, useful for

XX controlling carbon and nitrogen partitioning pathways during plant growth

XX and development.

PS Claim 18; Page 45; 48pp; English.

XX This partial cDNA encodes maize (Zea mays) sucrose non-fermenting 4 (SNF4) protein which is involved in carbon catabolite repression in plants and seeds. The cDNA was isolated, based on similarity to SNF4 proteins from Arabidopsis thaliana and Saccharomyces cerevisiae, from a library csk1c prepared from corn unpollinated developing silk 24 hours after emergence. The polynucleotides are used in plants to control carbon and nitrogen partitioning pathways during plant growth and development. The catabolite repression proteins would facilitate studies for better understanding the mechanism of catabolite repression in plants and could provide genetic tools to enhance or otherwise alter the accumulation of carbohydrates, lipids and proteins during plant growth and development

XX Sequence 538 BP; 146 A; 115 C; 120 G; 156 T; 0 U; 1 Other;

Alignment Scores: 1.68e-82 Length: 538

Pred. No.: 923.00 Matches: 174

Score: 923.00 Conservative: 2

Percent Similarity: 98.32% Mismatches: 3

Best Local Similarity: 97.21% Indels: 0

Query Match: 39.41% Gaps: 0

DB: 3

US-09-857-525C-2 (1-451) x AAAS1354 (1-538)

QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22

Db 2 GAACATTTGCCGATGTTCTCTATCGAAGCTGCCCACTGATTTTCAGGCTATTTCAGCT 61

QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu 42

Db 62 CTGCTCTCTGGGATTCAGAGTACAAATTCCTTTGGAGCGGGAATGGCGGATGATGAG 121

QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62

Db 122 CTTCAACCTACCATATCTGGGGATTTGGCATAGTTAACACACTTTACTTGACAAGGAA 181

QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82

Db 182 TATAACCAAAATAACACCTTATCAAGTCCAGCACACCTCGGAAGCAGGATGAACATGGAT 241

QY 83 ValAspAsnGlnAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102

Db 242 GTGGATAATGAATAATTTCAACGTACGGTTACGTTGTCAGATGCCACCGTTTCAGAGGT 301

QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122

Db 302 ACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTCGTGTTTCTGATATCTG 361

QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142

Db 362 AATTTGCATACATGCTATGATTTACTCCAGATTTCTGGCAAGGTTTATTCGCTTAGACAT 421

QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGlnGlnGlyIleProValAlaPro 162

Db 422 AATTTACCTGTGAAGCAATCAATTCATATTCCTCATGAACAGGGGATTTCTGTAGTCTCT 481

QY 163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIle 181

Db 482 CTCTGGGACTCANTCAAGGTCAATTTGGTGGGCCCTTAGCCCAATGGATTTCAATA 518

RESULT 7

AAAS1357

ID AAAS1357 standard; cDNA; 514 BP.

XX AC AAAS1357;

XX AC AAAS1357;

DT 26-SEP-2000 (first entry)

XX Wheat sucrose non-fermenting 4 protein cDNA from clone wreln.pk0143.e2.

XX Wheat; sucrose non-fermenting 4 protein; SNF4; plant; seed; growth;

CC hybridisation or amplification technologies, as query sequences to
CC determine homology of known sequences, as probe for use in Southern or
CC Northern hybridisation, and to identify the presence of and/or to
CC determine the degree of similarity between two (or more) nucleic acid
CC sequences

RESULT 10
ID AAL48614 standard; cDNA; 1167 BP.
XX AAL48614;
AC AAL48614;
XX 11-OCT-2002 (first entry)
XX Human insulin receptor signaling modifier cDNA SEQ ID NO: 11.
XX Human insulin receptor signaling; insulin receptor signaling modifier;
KW Human; insulin receptor signaling; insulin receptor signaling modifier;
KW ISM; diabetes; metabolic syndrome; antidiabetic; gene, ss.
XX Homo sapiens.
OS
XX WO200255664-A2.
EN
XX 18-JUL-2002.
PD
XX 11-JAN-2002; 2002WO-US001048.
PF
XX 12-JAN-2001; 2001US-0261226P.
PR
XX 12-JAN-2001; 2001US-0261303P.
PR
XX 12-JAN-2001; 2001US-0261304P.
PR
XX 12-JAN-2001; 2001US-0261335P.
PR
XX 12-JAN-2001; 2001US-0261361P.
PR
XX 12-JAN-2001; 2001US-0261361P.
PR
XX 12-JAN-2001; 2001US-0261456P.
PR
XX 12-JAN-2001; 2001US-0261457P.
PR
XX 12-JAN-2001; 2001US-0261458P.
PR
XX 12-JAN-2001; 2001US-0261459P.
PR
XX 12-JAN-2001; 2001US-0261461P.
PR
XX 12-JAN-2001; 2001US-0261518P.
PR
XX 12-JAN-2001; 2001US-0261531P.
PR
XX 12-JAN-2001; 2001US-0261532P.
PR
XX 12-JAN-2001; 2001US-0261589P.
PR
XX 12-JAN-2001; 2001US-0261590P.
PR
XX 12-JAN-2001; 2001US-0261694P.
PR
XX 12-JAN-2001; 2001US-0261695P.
PR
XX 12-JAN-2001; 2001US-0261697P.
XX
FA (EXEL-) EXELIXIS INC.
XX
XX Seidel-Dugan C, Ferguson KC, Kidd T;
XX WPI; 2002-599664/64.
XX P-PSDB; AAO18495.
XX
XX Identifying an insulin receptor signaling modulator, useful as drug
XX targets for treating diabetes or metabolic disorders, comprises
XX contacting an assay system comprising insulin receptor signaling
XX modifiers with a test agent.
XX
XX Disclosure; Page 48; 232pp; English.
XX
XX The present invention relates to a method of identifying a candidate
XX insulin receptor (INR) signaling modulating agent, involving contacting
XX an assay system comprising an insulin receptor signaling modifier (ISM)
XX polypeptide or nucleic acid with a test agent, and detecting a test agent
XX -biased activity of the assay system. The method is useful for
XX identifying candidate INR signaling modulating agents. ISM genes may be
XX used as drug targets for treatment of disorders related to INR signaling
XX such as diabetes or metabolic syndrome. ISM nucleic acids and
XX polypeptides are useful for identifying and testing agents that modulate
XX INR function, and for other applications related to the involvement of ISM
XX in INR signaling, and for identifying subjects having a predisposition to
XX such diseases associated with INR signaling. The present sequence is an
XX ISM coding sequence described in the exemplification of the invention
XX
XX Sequence 1167 BP; 352 A; 246 C; 265 G; 304 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 3.37e-27 Length: 1167

Score: 373.50 Matches: 112
Percent Similarity: 49.87% Conservative: 79
Best Local Similarity: 29.24% Mismatches: 130
Query Match: 15.95% Indels: 62
DB: 6 Gaps: 15
US-09-857-525c-2 (1-451) x AAL48614 (1-1167)
QY 70 SerSerProSerThrProGlySerArg-MetAsnMetAspValAspAsnGluAsnPheG1 89
Db 45 GCGAGCCCTGGGACCCGCGGAGCCGCGATCTGAGAGAGCTGGAGTTTCGAG----- 96
QY 89 nAargThrValLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGluAl 109
Db 97 -----GACGAGCAGTAGAAGAC-----TCAGAAG 122
QY 109 aAlaIleGlnIleSerArgCysArgValSerGluThrLeuAsnLeuHisThrCysTyrAs 129
Db 123 TGGTGTATTACATG-----CGATTTCATGAGGTACACAAAGTGTATTGA 164
QY 129 pLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsnLeuProValLysGlnSe 149
Db 165 CATCGTTCACACCCAGTTCAGAGCTTGTCTTTGATACATACATTAACAAGTTAAAGGC 224
QY 149 rPheHisIleLeuHisGluGlnGlyLeuProValAlaProLeuTrpAspSerPheArgG1 169
Db 225 CTCTCTTGGCTTGGTAGCCCAACGGTGTCCGAGCAGCGCCACCTGTGGAGAGTAAACAA 284
QY 169 yGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArgGluLeuGluTh 189
Db 285 AAGTTTGTAGGAATGCTAACCAATTACAGATTTCAATAATATATACATACATAGATAC 339
QY 189 rHisGlySerAsnLeuThrGlu---GlnGlnLeuGluThrHisThrIleSerAlaTrpLy 208
Db 340 -TATAATCACCTATGCTATGAGATTTTGAATGAGAGAACATAAATGAAACATGGAG 398
QY 208 sGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHisLe 228
Db 399 GGAGCTTTATTACAA-----GAAACATTTAAGCCT-----TT 431
QY 228 uValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeuLeuGlnAsnG1 248
Db 432 AGTGAATATATCTCCAGATGCAAGCCTTTCGATGCTGTATATCTCTTGATCAAAATAA 491
QY 248 yIleSerThrValProValIleTyrSerSerSerSerAspGlySerPheProGlnLeuLe 268
Db 492 AATCCACAGATTGCCCGTTATTGACCCCTATCAGTGGGAATGCA-----CT 536
QY 268 uHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPheLysAsnSerTh 288
Db 537 TTATACACTTACCACACAAAAGAAATCTCAAGTTCTCTC---CAGCTTTTATGTCTGATAT 593
QY 288 rGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpValPr 308
Db 594 GCMAAGCCTGCTTCATGAAGCAGAACCTGGAGAGCTTGGAAATAGGAACGATAC----- 648
QY 308 oLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSe 328
Db 649 -----CACAACTGCTTCATACATCCACACACATCCCATCAT 686
QY 328 rSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspPheAs 348
Db 687 CAAAGCCTTGAACATATTTGTGGAAGACGAATATCAGCTCTGCTGCTGTGGATGAGTC 746
QY 348 nAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeuAlaLysAspLysVa 368
Db 747 AGGAAAGTGTAGATATTATTCCAAATTTGAATTAATTAATCTGCTCTGAGAAAC 806
QY 368 lTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeuGlyGlnAs 388
Db 807 ATACAATACCTTA-----GATATCAGGTGACCCAGCCCTTCAGCACCGCTTCA 857
QY 388 pAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMetCysLeuArgSerAs 408


```

Db 689 GGAATAGGAGCTAC-----CACAACATTCCTTCATA 721
Qy 322 ArgProAnalaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341
Db 722 CATCCAGACATCCCATCATCAAGCCCTTGAACATATTGTGGAAAGCAGAAATATCAGCT 781
Qy 342 IleProIleValAspAsnAspSerLeuLeuAspThrThrLeuSerSerArgSerAspIleThr 361
Db 782 CTGCCTGTGTGGATGAGTCTAGGAAAGTGTGAGATATTATTCACAAATTTGATGTAATT 841
Qy 362 AlaLeuAlaLeuAspLeuValThrHisValArgLeuAspGluMetThrIleHisGln 381
Db 842 AATCTGTCTGCGAAGAAACATCAATACCTA-----GATATCAGGTGACCCAG 892
Qy 382 AlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys 401
Db 893 GCCCTTCAGCACCGCTTCACAG-----TATTTTGAAGGT-----GTT 928
Qy 402 GlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGly 421
Db 929 GTGAAGTGCATTAAGCTGGAATATCTGGAGACCCTGGGACAGATGATGATGAGCTGAG 988
Qy 422 ValArgArgValPheIleVal---GluAlaGlySerLysArgValGluGlyIleIleSer 440
Db 989 GTCCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1042
Qy 441 LeuSerAspIlePheLysPheLeuLeu 449
Db 1043 CTGTCGACATCTCGACCGCTGATC 1069

RESULT 12
AAH14839
ID AAH14839 standard; cDNA; 2223 BP.
XX
AC AAH14839;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12660.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000JP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
XX
PR 27-AUG-1999; 99JP-00300253.
XX
PR 11-JAN-2000; 2000JP-00118776.
XX
PR 02-MAY-2000; 2000JP-00183767.
XX
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
WP1; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 12660; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:

```

```

CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence and the 3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 2223 BP; 633 A; 431 C; 496 G; 663 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.:          Length:          2223
Score:              373.00           Matches: 104
Percent Similarity: 51.20%           Conservative: 75
Best Local Similarity: 29.80%         Mismatches: 128
Query Match:        15.93%           Indels: 42
DB:                 4                Gaps: 13

```

US-09-857-525C-2 (1-451) x AAH14839 (1-2223)

```

Qy 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgVal---SerGluTyrLeu 122
Db 172 CTGGAGTTCGAGGAGCAAGCAGTAGAAGACTCAGAAAGTGGTGTTCATGCGATTCATG 231
Qy 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
Db 232 AGCTCACACACAGCTTATGACATCGTTCACACAGTTCACAGCTTCTGCTTTGTACT 291
Qy 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
Db 292 ACATTACAAGTTAAAAAGGCTTCTTTGCTTTGGTAGCCACCGTTCGCGACGACGCCA 351
Qy 163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db 352 CTGTGGGAGAGTAAAAACAAAGTTTGTAGGAATGCTACATTAACATTCATAAAT 411
Qy 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGlu---GluGlnLeuGluThr 201
Db 412 ATACTACATAGATAC-----TATAAATCACCTATGTTACAGATTATGAATTAGAGGA 465
Qy 202 HisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln 221
Db 466 CATAAATTTGAACATCGAGGGAGCTTTATTACAA-----GAAACA 507
Qy 222 TrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAla 241
Db 508 TTTAAGCCT-----TTAGTGAATATATCTCCAGATGCAAGCCTCTTCGATGCTGTA 558
Qy 242 ValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAsp 261
Db 559 TACTCTCTTCATCAAAAATAAATCCACAGATTCGCGCTTATTGACCTTATCATGTGGAAAT 618
Qy 262 GlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCys 281
Db 619 GCA-----CTTTATATACATTACCAAAAAGAAATCCTCAAGTTCCTTC--- 660
Qy 282 ArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
Db 661 CAGCTTTTATGTCGATATGCCAAGCTGCTTCATGACGACAGAACCTGGATGAGCTT 720

```

	QY	302	ProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeu	321
	Dd	:	: :: :: :	::: :: :
	Db	721	GGAATAGGACGGTAC-----CACAACATTGCCTTCATA	753
	QY	322	ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer	341
	Dd	:	: :: :: :	:::~::~ :: :
	Db	754	CATCCAGACTCCCATCATCAAGCGCTTGACATATTTGTGGAAAGACCAATATCAGCT	813
	QY	342	IleProileValAspAspAsnAspSerLeuLeuAspThrTyrserArgSerAspile <thr< th=""></thr<>	361
	Dd	:	: :: :: :	:::~::~ :: :
	Db	814	CTGCCCTGTTGTGATGAGTCAGGAAAAGTTGTAGATATTTATTCCAAATTTGTAATT	873
	QY	362	AlaleualalaspysValThyThrHisvalAargLeuspGlumetThrIleHisGln	381
	Dd	:	: :: :: :	:::~::~ :: :
	Db	874	AATCTTGCTGCTGAGAANAACATACATACCCTA-----GATATCACGGTAGCCCCAG	924
	QY	382	AlalaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGinArgCys	401
	Dd	:	: :: :: :	:::~::~ :: :
	Db	925	GCCTTCAGCACCGTTCACAG-----TATTTTGAAGCT-----GTT	960
	QY	402	GlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGly	421
	Dd	:	: :: :: :	:::~::~ :: :
	Db	961	GTGAAGTGCAATAAGCTGGAAATACCTGGAGACCATCGTGACACAGATAGTAAGACGCTGAG	1020
	QY	422	ValArgArgValPheIleVal---GluAlaGlySerLysArgValGlnGlyVilIeSer	440
	Dd	:	: :: :: :	:::~::~ :: :
	Db	1021	GTCCATCGGCTGGTGGTGGAATGAAGACAGATAGT-----ATTGTGGGTATTATTTC	1074
	QY	441	LeuSerAspilePhelypSphLeuLeu	449
	Dd	:	: :: :: :	:::~::~ :: :
	Db	1075	CTGTGGACATTTGCAAGCGCTGATC	1101
		RESULT 13		
		AAL48615		
		ID AAL48615 standard; cDNA; 2223 BP.		
	Xx	AAL48615;		
	Xx	XX		
	Dt	11-OCT-2002 (first entry)		
	Xx	XX		
	De	Human insulin receptor signaling modifier cdna SEQ ID NO: 13.		
	Xx	XX		
	Kw	Human; insulin receptor signaling; insulin receptor signaling modifier;		
	Xx	XX		
	Os	ISM; diabetes; metabolic syndrome; antidiabetic; gene; ss.		
	Xx	Homo sapiens.		
	Pn	W0200255664-A2.		
	Xx	XX		
	Pd	19-JUL-2002.		
	Pf	XX		
	Xx	11-JAN-2002; 2002WO-US001048.		
	Pr	12-JAN-2001; 2001US-0261226P.		
	Pr	12-JAN-2001; 2001US-0261303P.		
	Pr	12-JAN-2001; 2001US-0261304P.		
	Pr	12-JAN-2001; 2001US-0261335P.		
	Pr	12-JAN-2001; 2001US-0261336P.		
	Pr	12-JAN-2001; 2001US-0261361P.		
	Pr	12-JAN-2001; 2001US-0261456P.		
	Pr	12-JAN-2001; 2001US-0261457P.		
	Pr	12-JAN-2001; 2001US-0261458P.		
	Pr	12-JAN-2001; 2001US-0261459P.		
	Pr	12-JAN-2001; 2001US-0261461P.		
	Pr	12-JAN-2001; 2001US-0261518P.		
	Pr	12-JAN-2001; 2001US-0261531P.		
	Pr	12-JAN-2001; 2001US-0261532P.		
	Pr	12-JAN-2001; 2001US-0261533P.		
	Pr	12-JAN-2001; 2001US-0261589P.		
	Pr	12-JAN-2001; 2001US-0261590P.		
	Pr	12-JAN-2001; 2001US-0261594P.		
	Pr	12-JAN-2001; 2001US-0261695P.		
	Xx	12-JAN-2001; 2001US-0261697P.		


```
Db 619 GCA-----CTTATATACCTTACCCCAAAAGATCCTCAAGTTCCTC--- 660
Qy 282 ArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIle 301
Db 661 CAGCTTTTATGTCGATATGCAAGAGCTGCTTTCATGAAGCAGAACTGGATGAGCTT 720
Qy 302 ProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProIleAlaMetLeu 321
Db 721 GGAATAGGAACGTAC-----CACAACTTGCCTTCATA 753
Qy 322 ArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSer 341
Db 754 CATCCAGACACTCCCATCATCAAGACCTTGAACATATTGTGGAAAGAGCAATATCAGCT 813
Qy 342 IleProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThr 361
Db 814 CTGCTGTGTGGATGAGTCAGAAAGTTGTAGATATTATTCCAAATTTGATGTAAAT 873
Qy 362 AlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGln 381
Db 874 AATCTTGTCTGAGAAACATACATAACCTA-----GATATCAGGTGACCCAG 924
Qy 382 AlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCys 401
Db 925 GCCCTTACGCCCTTCACAG-----TATTTGAAGT-----GTT 960
Qy 402 GlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGly 421
Db 961 GTGAAGTGCATTAAGCTGGAATACTGGAGACCATCGTGGACAGAATAGTAGAGCTGAG 1020
Qy 422 ValArgArgValPheIleVal---GluAlaGlySerLysArgValGluGlyIleIleSer 440
Db 1021 GTCATATCGCTGCTGCTGGTGAATGAAGCAGATAGT-----ATTGTGGTATTATTCC 1074
Qy 441 LeuSerAspIlePheLysPheLeuLeu 449
Db 1075 CTGTCGGACATCTGCAAGCCCTGATC 1101
RESULT 14
ID AAL48616 standard; cDNA; 1578 BP.
AC AAL48616;
XX
DT 11-OCT-2002 (first entry)
DE Human insulin receptor signaling modifier cDNA SEQ ID NO: 15.
KW Human; insulin receptor signaling; insulin receptor signaling modifier;
KW ISM; diabetes; metabolic syndrome; antidiabetic; gene; ss.
XX Homo sapiens.
XX WC2002055664-A2.
XX
PD 18-JUL-2002.
XX
PF 11-JAN-2002; 2002WO-US001048.
XX
PR 12-JAN-2001; 2001US-0261226P.
PR 12-JAN-2001; 2001US-0261303P.
PR 12-JAN-2001; 2001US-0261304P.
PR 12-JAN-2001; 2001US-0261335P.
PR 12-JAN-2001; 2001US-0261336P.
PR 12-JAN-2001; 2001US-0261361P.
PR 12-JAN-2001; 2001US-0261456P.
PR 12-JAN-2001; 2001US-0261457P.
PR 12-JAN-2001; 2001US-0261458P.
PR 12-JAN-2001; 2001US-0261459P.
PR 12-JAN-2001; 2001US-0261461P.
PR 12-JAN-2001; 2001US-0261518P.
PR 12-JAN-2001; 2001US-0261531P.
```

```
PR 12-JAN-2001; 2001US-0261532P.
PR 12-JAN-2001; 2001US-0261589P.
PR 12-JAN-2001; 2001US-0261590P.
PR 12-JAN-2001; 2001US-0261694P.
PR 12-JAN-2001; 2001US-0261695P.
PR 12-JAN-2001; 2001US-0261697P.
XX (EXEL-) EXELIXIS INC.
XX
PI Seidel-Dugan C, Ferguson KC, Kidd T;
XX
DR WPI: 2002-599564/64.
XX P-PSDB; AAO18497.
XX
PT Identifying an insulin receptor signaling modulator, useful as drug
PT targets for treating diabetes or metabolic disorders, comprises
PT contacting an assay system comprising insulin receptor signaling
PT modifiers with a test agent.
XX
PS Disclosure; Page 53-54; 232pp; English.
XX
CC The present invention relates to a method of identifying a candidate
CC insulin receptor (INR) signaling modulating agent, involving contacting
CC an assay system comprising an insulin receptor signaling modifier (ISM)
CC polypeptide or nucleic acid with a test agent, and detecting a test agent
CC biased activity of the assay system. The method is useful for
CC identifying candidate INR signaling modulating agents. ISM genes may be
CC used as drug targets for treatment of disorders related to INR signaling
CC such as diabetes or metabolic syndrome. ISM nucleic acids and
CC polypeptides are useful for identifying and testing agents that modulate
CC ISM function and for other applications related to the involvement of ISM
CC in INR signaling, and for identifying subjects having a predisposition to
CC such diseases associated with INR signaling. The present sequence is an
CC ISM coding sequence described in the exemplification of the invention
XX
SQ Sequence 1578 BP; 400 A; 377 C; 379 G; 422 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,348-26 Length: 1578
Score: 367.00 Matches: 111
Percent Similarity: 48.43% Conservative: 74
Best Local Similarity: 29.06% Mismatches: 133
Query Match: 15.67% Indels: 64
DB: 6 Gaps: 14
US-09-857-525C-2 (1-451) x AAL48616 (1-1578)
Qy 69 LeuSerSerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAsnPhe 88
Db 107 ATTTCTTCAGATAGCTCCCGAGCT-----GTGAAATAGCATCCT 148
Qy 89 GlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGlu 108
Db 149 CAAGAGACCCAGAA---TCCAAACAATAGCGTG----- 178
Qy 109 AlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTyr 128
Db 179 -----TATATCTTCATGAGTCTCATCGTCTAT 211
Qy 129 AspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsnLeuProValLysGln 148
Db 212 GACCTGATTCACCAAGCTCCAAATTTGGTTGATTGATACCTCCCTGCAGGTGAAGAA 271
Qy 149 SerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTyrAspSerPheArg 168
Db 272 GCTTTTTCCTTTCCTGACTAACCGTGTACGAGCTGCCCTTTATGGATAGTAGAAG 331
Qy 169 GlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuLeuArgGluLeuGlu 188
Db 332 CAAAGTTTTTGGGATGCTGACCATCATGATTTTCATCAATATCTTGCACCGCTAC--- 388
Qy 189 ThrHisGlySerAsnLeuThrGlu---GluGlnLeuGluThrHisThrIleSerAlaTrp 207
```

389	Db	---	TATTAATCAGCCCTTGGTACAGATCTATGAGCTAGAGAACACACAGATGAACAACCTTGG	445
208	Qy	Lys	GluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnHis	227
446	Db	AGAGAGGTGATCTCCAG-----	-GACTCCTTTAAACCG-----	478
228	Qy	Leu	ValHisalaThrProTyrGluSerLeuArgAspIleAlaValLysLeuLeuGlnAsn	247
479	Db	CTTGCTCGCATTTCTCCTAATGCCAGCTGTCTTGATGCTGCTCTTCATTAATTTCGGAAC	538	
248	Qy	GlyIleSerThrValProValIleTyrSerSerSerSerAspGlySerPheProGlnLeu	267	
539	Db	AAGATCCACAGGCTGCCAGTATTGACCCAGAACTAGGCAT-----	-ACT	583
268	Qy	LeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPheLysAsnSer	287	
584	Db	TTGTACATCTCCACCCCAACGCGATTCTGAAGTTTCCTC---AAATGTTTATCACTGAG	640	
288	Qy	ThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpVal	307	
641	Db	TTCCCCAAGCCAGGTTCACTGTCCAGTCTCTGGAAGAGCTACAGATTGCNCTATGCC	700	
308	Qy	ProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeu	327	
701	Db	-----AATATTGCTATGGTTTCGCACTACCAACCCCGCTC	733	
328	Qy	SerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspAsp	347	
734	Db	TATGTGCTCTGGGGATTTTGTACAGATCGAGTCTCAGCCCTGCCAGTGGTGATGAG	793	
348	Qy	AsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeuAlaLysAspLys	367	
794	Db	AAGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTTATCAATCTGGCAGCAGAAAG	853	
368	Qy	ValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeuGlyGln	387	
854	Db	ACCTACAAACAACCTA-----GATGTATCTGTGACTAAAGGCTTGCACATCGATCA	904	
388	Qy	AspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMetCysLeuArgSer	407	
905	Db	CAT-----TACTTTGAGGGT-----GTTCTCAAGTGTACTCTGGAT	940	
408	Qy	AspProLeuLeuLysValMetGluArgGluAlaAsnProGlyValArgValPheIle	427	
941	Db	GAGACTCTGGAGACCATCATCAACAGGCTAGTGAAGCAGAGG-TCCACCGCATGTGATG	1000	
428	Qy	ValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePheLysPhe	447	
1001	Db	GTGATGAAATGATGTG-----GTCAAGGGAATGTATCACTGTCTGACATCTCTCAGGCC	1057	
448	Qy	LeuLeu	449	
1058	Db	CTGGTG	1063	

RESULT 15

ABK84324

ID ABK84324 standard; cDNA; 1578 BP.

XX

AC ABK84324;

DT 14-AUG-2002 (first entry)

XX DE Human cDNA differentially expressed in granulocytic cells #895.

Human; ss; granulocytic cell; DNA chip; bacterial infection; viral infection; parasitic infection; protozoal infection; fungal infection; sterile inflammatory disease; psoriasis; rheumatoid arthritis; glomerulonephritis; asthma; thrombosis; cardiac reperfusion injury; renal reperfusion injury; ARDS; adult respiratory distress syndrome; inflammatory bowel disease; Crohn's disease; ulcerative colitis; periodontal disease; granulocyte activation; chronic inflammation; allergy.

OS Homo sapiens.
XX WO200228999-A2.
XX 11-APR-2002.
XX 03-OCT-2001; 2001WO-US030821.
XX 03-OCT-2000; 2000US-0237189P.
XX (GENE-) GENE LOGIC INC.
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
XX Detecting granulocyte activation by detecting differential expression of
XX genes associated with granulocyte activation, which serves as diagnostic
XX markers that is useful for monitoring disease states and drug toxicity.
XX
XX Claim 1: SEQ ID NO 395; 114pp; English.
XX
XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing the
XX expression level to an expression level in an unactivated GC, where
XX differential expression of Gs is indicative of GCA. Also included are
XX modulating (M2) GA by contacting GC with an agent that alters the
XX expression of at least one gene in Gs; (2) screening (M3) for an agent
XX capable of modulating GCA or an inflammation (especially chronic) in a
XX tissue, an allergic response in a subject, exposure of a subject to a
XX pathogen or sterile inflammatory disease using the gene expression
XX profile; (3) detecting (M4) an inflammation (especially chronic) in a
XX tissue, an allergic response in a subject, exposure of a subject to a
XX pathogen or sterile inflammatory disease, by detecting the level of
XX expression in a sample of the tissue of gene(s) from Gs, where the level
XX of expression of the gene is indicative of inflammation; (4) treating
XX (M5) an inflammation (especially chronic) or in a tissue, an allergic
XX response in a subject, exposure of a subject to a pathogen or sterile
XX inflammatory disease, by contacting a tissue having inflammation with an
XX agent that modulates the expression of gene(s) from Gs in the tissue. M1
XX is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
XX for screening an agent capable of modulating GCA preferably in an
XX inflammation in a tissue; M4 is useful for detecting an inflammation
XX (especially chronic) in a tissue, an allergic response in a subject,
XX exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
XX psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
XX cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
XX respiratory distress syndrome, inflammatory bowel disease, Crohn's
XX disease, ulcerative colitis, periodontal disease; also bacterial
XX infection, viral infection, parasitic infection, protozoal infection,
XX fungal infection and M5 is useful for treating one of the above
XX conditions. The present sequence represents a gene differentially
XX expressed in granulocytes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

Alignment Scores:

Pred. No.:	2.34e-26	Length:	1578
Score:	367.00	Matches:	111
Percent Similarity:	48.43%	Conservative:	74
Best Local Similarity:	29.06%	Mismatches:	133
Query Match:	15.67%	Indels:	64
DB:	5	Gaps:	14

US-09-857-525C-2 (1-451) x ABK84324 (1-1578)

Qy 69 LeuSerSerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAsnPro 88
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
107 ATTCTTCAGATAGCTCCCACT-----GTGAAATGACATCACT 148
DB

```
Qy      89  GluArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGlu 108
Db      149  CAAGAGACCCAGAA---TCCACATAGCGT-----178
Qy      109  AlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTyr 128
Db      179  -----TATACCTCCTCATGAAGTCTCATCGCTGCTAT 211
Qy      129  AspLeuLeuProAspSerGlyValIleAlaLeuAspIleAsnLeuProValIysGln 148
Db      212  GACCTGATCCACAGCTCCAAATGGTTGATTGATAGCTCCCTGCAGGTGAAGAAA 271
Qy      149  SerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuThrAspSerPheArg 168
Db      272  SCITTTTGTCTTTGGTGACTAACGGGTGTACAGCTGCCCTTTATGGGATAGTAAGAAG 331
Qy      169  GlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArgGluLeuGlu 188
Db      332  CAAAGTTTGTGGGATGCTGACCATCTACTGATTTTCATCAATNTCCCTGCACCGCTAC--- 388
Qy      189  ThrHisGlySerAsnLeuThrGlu---GluGlnLeuGluThrHisThrIleSerAlaTrp 207
Db      389  ---TATAAATCAGCCTTGGTACAGATCTATGAGCTAGAGAACACACAGATAGAAACTTGG 445
Qy      208  LysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHis 227
Db      446  AGAGAGGTGTATCTCCAG-----GACTCCTTTAAACOG----- 478
Qy      228  LeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValIysLeuLeuGlnAsn 247
Db      479  CTGTCTGCATTTCTCCAAAGCCAGCTTGTGATGCTGCTCTCTTCATTAATTCGGAC 538
Qy      248  GlyIleSerThrValProValIleTyrSerSerSerAspGlySerPheProGlnLeu 267
Db      539  AAGATCCACAGGCTGCCAGTTATTGACCCAGAAATCAGGCAAT-----ACT 583
Qy      268  LeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPheLysAsnSer 287
Db      584  TTGATACATCTCCACCAAGCGCATCTCTGAAGTTCTTC---AAATGTTTTATCACTGAG 640
Qy      288  ThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpVal 307
Db      641  TTCCCCAAGCCAGAGTTCATGTCCTCAGGCTCTGGAGAGCTACAGATTGGCACCTATGCC 700
Qy      308  ProLysIleGlyAspLeuAsnSerArgProLeuAlaValMetLeuArgProAsnAlaSerLeu 327
Db      701  -----AATATTGCTATGTTGCACTACCAACCCCGCTC 733
Qy      328  SerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspAsp 347
Db      734  TATGTGGCTCTGGGGATTTTGTACAGCATCGAGTCTCAGCCCTCCAGTGGTGGATGAG 793
Qy      348  AsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeuAlaLysAspLys 367
Db      794  AAGGGGCGTGTGGTGACATCTACTCCAAGTTGATTTATCAATCTGGCAGCAGAAAAAG 853
Qy      368  ValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeuGln 387
Db      854  ACCTACAAACACCTA-----GATGTATCTGTGACTAAAGCCTTGCAACATCGATCA 904
Qy      388  AspAlaAsnThrProPhePheAsnGlyGlnArgCysGlnMetCysLeuArgSer 407
Db      905  CAT-----TACTTCAGGGT-----GTTCTCAAGTGTACCTGCAT 940
Qy      408  AspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyValArgValPheIle 427
Db      941  GAGACTCTGGAGACCATCATCAACAGGTAGTGGAGCAGAGGTTCAACCGACTGTAGTG 1000
Qy      428  ValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePheLysPhe 447
Db      1001  GTGGATGAAATGATGTG---GTCAAGGGAAATTGTATCACTGCTGACATCTCTGCAGGCC 1057
```

Qy 448 LeuLeu 449
Db 1058 CTGGTG 1063

Search completed: July 9, 2004, 08:38:22
Job time : 628 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OX protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2004, 07:01:49 ; Search time 3645 Seconds
(without alignments)
3694.882 Million cell updates/sec

Title: US-09-857-525C-2
Perfect score: 2342
Sequence: 1 TREHLPSPBPGCTVQAL.....SKRVEGIISLSIDIFKFLSL 451

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09857525/runat_07072004_161018_305/app_query.fasta_1.647
-DB=EST -QMT=fastcap -SUPFI=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRP=US09857525 @CNG 1.1 3437 @runat_07072004_161018_305 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_yrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2107	90.0	2227	11	AY109512	Zea mays
2	1274	54.4	873	14	CD434895	EL01N0330
3	1090	46.5	666	14	CA211312	SCRLAD109
4	1087	46.4	644	10	AW744961	LGI_385_D
5	1086	46.4	738	14	CD878389	WHE4018_C
6	1064	45.4	737	14	CA500333	WHE4018_C
7	1050	44.8	712	14	CD903099	WHE4018_C
8	1045	44.6	683	14	CA262684	SCPLB202
9	1018	43.5	601	13	BU098591	SCPLB202
10	999	42.7	595	10	BE598304	PII_68_G0
11	969	41.4	577	14	CA218557	SCUFADIC0
12	964	41.2	559	10	BE123279	945040D06
13	963	41.1	549	13	BU499410	946174G09
14	953	40.7	631	14	CA253676	SCRLFLA10
15	943	40.3	831	10	BE704511	SC01_08c1
16	939	40.1	558	10	BE598914	PII_84_H0
17	913	39.0	726	14	CD902160	G356_106C
18	904	38.6	594	12	BU450369	BU450369
19	898	38.3	604	12	BU477917	BU477917
20	892	38.1	690	13	BQ163922	952081C10
21	891	38.0	534	12	BI423527	949049B31
22	889	38.0	563	12	BG605114	WHE2327_E
23	889	38.0	680	13	CA148077	SCERZ101
24	875	37.4	911	13	CA067477	SCQAD106
25	867	37.0	626	14	CA498139	WHE3239_C
26	855	36.5	612	14	CA595438	wpalc.p40
27	848.5	36.2	596	14	CF052805	QCM8C01_Y
28	838	35.8	689	14	CF635900	ZMTRW00_0
29	835	35.7	660	14	CB008818	VVC060FT1
30	828	35.4	693	14	CD863501	AZ01.106P
31	824	35.2	571	14	CF302989	ABFI_01-
32	814.5	34.8	745	14	CF438706	EST675051
33	814	34.8	700	13	CA124616	SCQGLR108
34	810.5	34.6	707	14	CF087439	OHM14P15
35	809	34.5	549	13	BU969888	HE12P01r
36	806	34.4	607	14	CK103865	1002P24.5
37	800	34.2	541	12	BJ303942	BJ303942
38	794.5	33.9	626	14	CD995954	QBB7C06_X
39	793.5	33.9	780	12	BI930509	EST550398
40	787	33.6	555	12	BI479562	WHE3455_B
41	785	33.5	530	14	CF325983	JMT11_04-
42	776.5	33.2	716	13	BU025775	QHG1IF23
43	769	32.8	712	10	BF625365	FVSNM000
44	767	32.7	473	10	AW676933	DGI_2_C07
45	767	32.7	524	12	BG560098	RHI22_69

ALIGNMENTS

RESULT 1
AY109512
LOCUS AY109512 2227 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL367_1 mRNA sequence.
ACCESSION AY109512
VERSION AY109512.1 GI:21213260
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2227)

AUTHORS Hainey,C.F., Dolan,M., Xiao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo probes

JOURNAL Unpublished (2002)

AUTHORS 2 (bases 1 to 2227)

TITLE Coe,E.H.

JOURNAL Direct Submission

COMMENT Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES Location/Qualifiers

source 1..2227

organism="Zea mays"

mol_type="mRNA"

db_xref="Xref:4577"

clone_lib="Maize Mapping Project/DuPont Consensus
Library"

note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Alignment Scores:

Pred. No.: 2,47e-236 Length: 2227

Score: 2107.00 Matches: 4:2

Percent Similarity: 92.87% Conservative: 5

Best Local Similarity: 91.76% Mismatches: 32

Query Match: 89.97% Indels: 0

DB: 11 Gaps: 0

US-09-857-525c-2 (1-451) x AY109512 (1-2227)

QY 3 GluHisLeuProMetSerProFleGluGlyCysProThrValPheGlnAlaIleCysSer 22
|||||
DB 640 GAACATTTGCCGATGCTCTCTGTCGAGGCTGCCCTGATTTTCAGGCTATTTCAGGC 699
|||||

QY 23 LeuSerProGlyIleHisGluTyrLysPheValAspGlyGluTyrArgHisAspGlu 42
|||||
DB 700 CTGCTCTCCAGGATTCATGATGATACAGTTCTATGTGGACGGGAGTGGGCTCATGATGAG 759
|||||

QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrIleThrArgGlu 62
|||||
DB 760 CGCAACCTACTATATCTCGGGAGTTCGGTATAGTTAAACACATATCTTGCACAGGGAA 819
|||||

QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
|||||
DB 820 TTTTAATCAAAATAATGCTCTTATTAATCAAGCACACCTGGAGCGAGTAACATGAT 879
|||||

QY 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
|||||
DB 880 GTGGATAATGAAATTTTCAAGCTACGTACGTTCAGTTGTCGATGTTACCGTTTCAGAGGT 939
|||||

QY 103 ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
|||||
DB 940 ACTCTGAGAGTTTCAGAGGCTGCAATACAAATATCTAGGTGTGCTGTTTCTGAGTATCTG 999
|||||

QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
|||||
DB 1000 AATTGCGNACATGCTATGATTCTCCCGGATTCGGCAAGGTATTGCGCTCGACATT 1059
|||||

QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
|||||

DB 1060 AATTACTCTGTGAAGCAATCATTTCCATATTTCCATGAACACGGGATTCCTGTAGTCCT 1119
QY 163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
|||||
DB 1120 CTCTGGGACTCATTCAGAGTCAATTTGTGGCTCTTCGAGCCCATTTGATTTTACATCT 1179
|||||

QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
|||||
DB 1180 ATATTTCGGGAGCTAGAACTCATGCTCGAACTTTGACAGAAGATCAGCTTGAACACAC 1239
|||||

QY 203 ThrIleSerAlaTyrLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTyr 222
|||||
DB 1240 ACTATATCTGCATGGAAGAGAGCTAAGCGGAACTTTGGAGAAATGATGTCTAGTGG 1299
|||||

QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
|||||
DB 1300 CGAGCACATCAGCATCTAGTGCATGCCACCTTATGAGTCTCTGAGGACATTCAGTA 1359
|||||

QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly 262
|||||
DB 1360 AAGCTTTTGTCTAAATGACATTTCTACAGTGCAGTATTTATTCATCATCAGATGGA 1419
|||||

QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
|||||
DB 1420 TCATTCCCTCAGTATTGACCTTCATCCCTTTCTGGAATTTTNNNNNNNNNNNNNN 1479
|||||

QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
|||||
DB 1480 NNNNNNNNNNTCACTGGTAATTTGCTATTTCTGAACCAACCGGTGTGCTCCATTCGG 1539
|||||

QY 303 LeuGlySerTyrValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 222
|||||
DB 1540 CTGGGTTCTCGGTTCCGAAATCGGTATCAAAACAGTCTGTCATTTGCTATGTTGCGA 1599
|||||

QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
|||||
DB 1600 CCTAATGTCATCACTTAGCTCTGCCCTTACATGTTGTTCAAGCTGGAGTGAGTCAATA 1659
|||||

QY 343 ProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
|||||
DB 1660 CCNNNNNNNNNNNNNNNNNNCTGCTTGCATCTTACTTAGAAGTGACATCAGCC 1719
|||||

QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
|||||
DB 1720 CTAGCTAAAGCAAGCTCTACACATGTTGCCCTGGATGAGATGCCATTCATCAGCT 1779
|||||

QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
|||||
DB 1780 TTACAGCTTGCACAAAGATGCCAATACGCTTTTGGANNNNNNAACGTCAGAGATGCCAG 1839
|||||

QY 403 MetCysLeuArgSerAspProLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
|||||
DB 1840 ATGTGCTCCGCTCGATCTCTTGTCTGAAGTGATGGAGCGCATCGCTAATCCNNNNNG 1899
|||||

QY 423 ArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSer 442
|||||
DB 1900 CGCGGGTGTTCATAGTAGAGCTGGAGCAAACTGTGGAGGGTATTATATCATCTAGT 1959
|||||

QY 443 AspIlePheLysPheLeuLeuSerLeu 451
|||||
DB 1960 GATATTTTCAAGTCTTCTGCTGAGCTTG 1986
|||||

RESULT 2

CD434895 873 bp mRNA linear EST 03-JUN-2003

LOCUS E010330A04.b EndospERM_3 Zea mays cDNA, mRNA sequence.

DEFINITION CD434895

ACCESSION CD434895

VERSION CD434895.1 GI:31350538

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Zea mays

ORGAN-SM

```

618 ACTTATATTGCGGGAGCTAGAAACTCATGGCTCGAACTTTGACAGAAGATCAGCTTGAAC 677
201 rhtsthrlleSerAlaTrrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerG1 221
678 ACACACTATATCTGCAATGAAAGAGGCTAAGCGGCAAACTTTGGAAGAAATATGATGTC 737
221 ntTrrArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAl 241
738 GTGGGGAGCACATCAGCACTAGTCATGCCACCCCTTATGAGTCCTTTGAGGGACATTGC 797
241 aValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAs 261
798 AGGTAGCTTTTGCAAAATGACATTTCTACAGTGCAGTCCAGTTATTTATTCATCATCAGA 857
261 pGlySerPheProGln 266
858 TGGATCATTCCTCCAG 873

RESULT 3
CA211312
LOCUS
DEFINITION
CA211312 666 bp mRNA linear EST 25-SEP-2003
SRLADI098B08.g Ad1 Saccharum officinarum cDNA clone SCLADI098B08
5', mRNA sequence.
CA211312
CA211312.1 GI:35255493
EST.
Saccharum officinarum
Saccharum officinarum
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 666)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 098 row: B column: 08
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .666
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SRLADI098B08"
/lab_host="DH103"
/clone_lib="AD1"
/note="Organ: seedlings inoculated with Gluconacetobacter
diazotrophicans; Vector: pSport1; Site_1: Sail; Site_2:
NotI; An unidirectional cDNA library generated from
seedlings inoculated with Gluconacetobacter
diazotrophicans). cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"

FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 2.43e-117 Length: 666
Score: 1090.00 Matches: 211
Percent Similarity: 95.5% Conservative: 2

```


Best Local Similarity: 94.62% Mismatches: 8
 Query Match: 46.54% Indels: 2
 DB: 14 Gaps: 0

US-09-857-525C-2 (1-45) x CA211312 (1-666)

```

QY 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerLeu 23
DB 1 CAATTTGCCGATGTCCTCTGTGCGAGGCTGCCCATCTGATTTTCAGGCTATTTGCGACCTG 60
QY 24 SerProGlyIleHisGluThrLysPhePheValAspGlyGluTrpArgHisAspGluArg 43
DB 61 TCTCCAGGATCCAGAGTACAGATTTTGTGGACGGGAGTGGCGCATGATGAGCGC 120
QY 44 GlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuThrArgGluThr 63
DB 121 CAACCTACTATATCCGGGAGTGTGGTATAGTTAACAACATCTACTTGCACAGAAATTT 180
QY 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83
DB 181 AACCAAAATAAACGCTTTATTAAGTCCAGCAAAATTTGGAAGCAGGATGAACATGATGTG 240
QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
DB 241 GATATGAAATTTTCAAGTACGTTAGTGTGCGATGGTACTGTTTCGGAGGTACT 300
QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluThrLeuAsn 123
DB 301 CCGAGAGTTTCAGAGGCTGCATACAAATCTCTAGTGCGCGCTTTCTGAATATCTGAAT 360
QY 124 LeuHisThrCysTrpAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
DB 361 TTGCATACATGATGATTTTCTCCGAGTTCAGGCAAGTTATTTCTCTGACATTAAT 420
QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
DB 421 TTACCTGTGAAGCAATCTTTTTCATATCTTACATGAACAGGGATTCCTGTAGCTCTCTG 480
QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183
DB 481 TGGGACTCATTCAGAGGTCAATTTGTGTCTCTCTGAGCCCATGAGTTTCATCTTATA 540
QY 184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnGlnLeuGluThrHisThr 203
DB 541 TCGCGGAGCTAGAACTCATCTGCTCGAACCTGCAGAGACAGCTTGAACACACACT 600
QY 204 TieSerAlaTrpLysGluAlaLysArgGlnThrAsnGly-ArgAsnAspSerGln-TrpA 223
DB 601 ATATCTGCATGGAAGAGGCTAACCGGCAAACTTTATCGAGAAATGATGTGTCATTTGCG 660
QY 223 r3Pro 224
DB 661 GACCG 665

```

RESULT 4
 AW744961
 LOCUS
 DEFINITION LG1_385.D10.bl_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
 sequence.
 ACCESSION AW744961
 VERSION AW744961.1 GI:7658699
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 644)
 Cordonnier-Pratt M.-M., Gingle, A., Marsala, C. and Pratt, L. H.
 An EST database from Sorghum: light-grown seedlings
 TITLE Unpublished (2000)
 JOURNAL Contact: Cordonnier-Pratt MM
 COMMENT Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: JEN REV
 High quality sequence stop: 621
 POLYA=No.

FEATURES

source

Location/Qualifiers
 1..644
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /db_xref="taxon:4558"
 /clone_lib="Light Grown 1 (LG1)"
 /note="Organ: 10- to 14-day-old light-grown (greenhouse)
 seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2:
 EcoRI; The library was made from poly-A RNA in the cloning
 vector lambda Zap II. Clones to be sequenced were
 prepared by mass excision."

ORIGIN

Alignment Scores:
 Pred. No.: 5,16e-117 Length: 644
 Score: 1087.00 Matches: 285
 Percent Similarity: 98.11% Conservative: 3
 Best Local Similarity: 96.70% Mismatches: 4
 Query Match: 46.41% Indels: 0
 DB: 10 Gaps: 0
 US-09-857-525C-2 (1-451) x AW744961 (1-644)
 QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 DB 4 GAACATTTGCCGATGTCCTCTGCGAGGCTGCCCATCTGATTTTCAGGCTATTTGCGAC 63
 QY 23 LeuSerProGlyIleHisGluThrLysPheValAspGlyGluTrpArgHisAspGlu 42
 DB 64 CTGCTCCAGGATTCAGAGTACAAAGTTCTTTTGGATGGGAGTGGCGCATGATGAG 123
 QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuThrArgGlu 62
 DB 124 CCGCAACCTACTATATCTGGGAGGTTTGGATAGTTTACACATTATCTTGACACGGGA 183
 QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 DB 184 TTTAACCACTAAACGCCCTTATTAAGTCCAGCACACCTTGGAGCAGGATGAACATGGAT 243
 QY 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 DB 244 GTGGATAATGAAAAATTTTCAAGTACGGTTAGTGTGCGATGGTACCGTTTCGGAGGT 303
 QY 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluThrLeu 122
 DB 304 ACTCCGAGAGTTTCAGAGGCTCAATACAAATCTCTAGTGCGCGCTTCTGAATATCTG 363
 QY 123 AsnLeuHisThrCysTrpAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
 DB 364 AATTTGSCATACATGCTATGATTTACTCCGGGATTCGGGCAAGGTTATTTGCTCTGACAT 423
 QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 DB 424 AATTTACTGTGAAGCAATCTTTTCATATCTGCAATGACAGGGATTTCTGTAGTCTCT 483
 QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
 DB 484 CTCTCGGACTCATTTAGAGGTCAATTTGTGTCTTCTGAGCCCATTTGGATTTCATCTT 543
 QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnGlnLeuGluThrHis 202
 DB 544 ATATTGGGGAGCTAGAAAACTCATGGCTCGAACTTTGACAGAGAGCAGCTTGAAACACAT 603

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

QY 203 ThrileSerAlaTrpLysGluAlaLysArgGlnThr 214
 DB 604 ACGATATCTGCATGGAAGAAGAGCTTAAGCGGCAAACT 639

RESULT 5
 CD878389 738 bp mRNA linear EST 11-JUL-2003
 LOCUS AZ04.102K03F010930 AZ04 Triticum aestivum cDNA clone AZ04102K03,
 DEFINITION mRNA sequence.
 ACCESSION CD878389
 VERSION CD878389.1 GI:32562205
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 738)
 AUTHORS Genoplante.
 TITLE Genoplante, a major partnership french program in plant genomics
 JOURNAL Unpublished (2003)
 COMMENT Contact: Genoplante
 Genoplante
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' {http://www.genoplante.com
 and http://genoplante-info.inbioigen.fr}.

FEATURES
 source
 1..738
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="recital"
 /db_xref="taxon:4565"
 /clone="AZ04102K03"
 /tissue_type="root"
 /clone_lib="AZ04"

ORIGIN
 Alignment Scores:
 Pred. No.: 8.71e-117 Length: 738
 Score: 1086.00 Matches: 209
 Percent Similarity: 91.80% Conservative: 15
 Best Local Similarity: 85.66% Mismatches: 18
 Query Match: 46.37% Indels: 2
 DB: 14 Gaps: 1

US-09-857-525C-2 (1-451) x CD878389 (1-738)

QY 75 ProGlySerArgMetAsnMetAspValAspAsnGluAsnPheGlnArgThrValThrLeu 94
 DB 12 CCTGGAGACAG-----ATGGATGTGGACAATGATAGTTTCAACGAATGGGTGCTG 65

QY 95 SerAspGlyThrValSerGluGlyThrLeuArgValSerGluAlaAlaIleGlnIleSer 114
 DB 66 TTGGATGTGTCCTTCAGGAAGGTCTCCAGAGATTTTCAGAGGCTGCTATACAGATCTCT 125

QY 115 ArgCysArgValSerGluThrLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSer 134
 DB 126 AGGTGTCTGTGTGTGATGATCTCAATGCGCATACAGGCCATGACCTACTACCAATCTCT 185

QY 135 GlyIysValIleAlaLeuAspIleAsnLeuProValIysGlnSerPheHisIleuHis 154
 DB 186 GGAAGATCATGCTCTGGACATTAATTTACCTGTGAAGCAACTTTTCATATCTCTCAT 245

QY 155 GluGlnGlyIleProValAlaProLeuTipAspSerPheArgGlyGlnPheValGlyLeu 174
 DB 246 GAACAGGGGATCTCTGTGGCTCTCTGTGGATTCATTCAGGGGTGAGTTGTGGCCTT 305

QY 175 LeuSerProLeuAspPheIleLeuIleLeuArgGluLeuGluThrHisGlySerAsnLeu 194

DB 306 CTGAGCCCACTGGATTTTATATCTATATTGAGAGAGCTGGAAACGCATGSGCTCAACCTG 365

QY 195 ThrGlnGlnLeuGluThrHisThrIleSerAlaTrpLysGluAlaLysArgGlnThr 214
 DB 366 ACAGAGAACAGCTTGAACAACACACATATATCTCGTGGAAAGAGGCTTAAGCGGCAAACT 425

QY 215 AsnGlyArgAsnAspSerGlnTipArgProGlnGlnHisLeuValHisAlaThrProTyr 234
 DB 426 TATGGAGAAATGATGGACAACCTTAGATCAATCAGCATCTAGTGCATGCCACCCCTTAT 485

QY 235 GluSerLeuArgAspIleAlaValIysLeuLeuGlnAsnGlyIleSerThrValProVal 254
 DB 486 GAATCCTTGGAGGCTTAATGCCATGAAATATCTTGAACCTGGTATTTCTACAGTACCAATC 545

QY 255 IleTyrSerSerSerSerAspGlySerPheProGlnGlnLeuHisLeuAlaSerLeuSer 274
 DB 546 ATCTATTCATCGTCATCAGATGGATCGTTTCGACGCTGTGCATCTTGATCCCTTCA 605

QY 275 GlyIleLeuLysCysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeu 294
 DB 606 GGAATTTTGAATGTATCTGTAGATACCTTAAGAACTCCACTGTGTAGTTTGGCATTTCTA 665

QY 295 AsnGlnProValCysSerIleProLeuGlySerTyrValProLysIleGlyAspLeuAsn 314
 DB 666 AACCAACCAAGTATGCTCAATTCGCTGGGTATACATGGGTTCACAAAATTTGGTGAACCAAT 725

QY 315 SerArgProLeu 318

DB 726 GGTATCTCATTTG 737

RESULT 6
 CA500333 737 bp mRNA linear EST 14-NOV-2002
 LOCUS WHE4018_C01_E02ZT Wheat meiotic anther cDNA library Triticum
 DEFINITION aestivum cDNA clone WHE4018_C01_E02, mRNA sequence.
 ACCESSION CA500333
 VERSION CA500333.1 GI:24991293
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatozoa; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 AUTHORS 1 (bases 1 to 737)
 Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R.,
 Pham, J., Rausch, C.J., Sutton, T., Woo, J., and Wilson, C.
 TITLE The structure and function of the expressed portion of the wheat
 genomes - Meiotic anther cDNA library
 JOURNAL Unpublished (2002)
 COMMENT Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105959773
 Fax: 5105959818
 Email: oanderson@pw.usda.gov
 Sequences have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: T7 primer.
 Location/Qualifiers
 1..737
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="WHE4018_C01_E02"
 /tissue_type="Anther"
 /dev_stage="Meiotic stages pre-meiosis-metaphase I"
 /lab_host="E. coli DH10B"
 /clone_lib="Wheat meiotic anther cDNA library"
 /note="vector: pSPOR1; Site1: SalI; Site2: NotI; Plants
 were grown in a glasshouse. Anther meiotic stage was
 determined by removing anthers from individual primary

florists. One author was sacrificed for microscopic staging, and if determined to be between (and including) meiotic stages pre-meiosis and metaphase I, the remaining two authors were collected and pooled for library construction. The tissue, total RNA, and poly(A) RNA were prepared, cDNA synthesized, and directionally ligated into pSPORT1 by Jim Sutton in the p Langridge lab at the Department of Plant Science, University of Adelaide, Waite Campus, Australia. Average insert size 1.5kb. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN

Alignment Scores:
 Pred. No.: 3,42e-114 Length: 737
 Score: 1064.00 Matches: 205
 Percent Similarity: 92.18% Conservative: 19
 Best Local Similarity: 84.36% Mismatches: 17
 Query Match: 45.43% Indels: 2
 DB: 14 Gaps: 1

US-09-857-525C-2 (1-451) x CA500333 (1-737)

QY 51 PheGlyIleValAsnThrLeuTyrLeuThrArgGluTyrAsnGlnIleAsnThrLeuSer 70
 Db 13 TATGGGGTGGTAACACCTTGTACTTGACAGAGGAATTGACCATAAATCTACTGTG 72
 QY 71 SerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAsnPheGlnArg 90
 Db 73 AGCCCACTACACCTGGGAGCAGG-----ATGGATGTGGACAATGATAGTTTCAACGA 126
 QY 91 ThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGluAla 110
 Db 127 ATGGGTTCCTGTTGGATGGTGGCCCTTCAGGAAGGTTCTCCAAGAATTTCAAGGCTCT 186
 QY 111 IleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTyrAspLeu 130
 Db 187 ATACAGATCTCTAGGTGTGGTGTCTGAGTATCTGGAATGGCATACAGGCGCATGACCTA 246
 QY 131 LeuProAspSerGlyValIleAlaLeuAspIleAsnLeuProValIysGlnSerPhe 150
 Db 247 CTACCAAGATCTGGNAAGCTCATGTCTGGACATTAATTTACCTGTGAGCAATCTTTC 306
 QY 151 HisIleLeuHisGluGlnGlyIleProValAlaProLeuTyrAspSerPheArgGlyGln 170
 Db 307 CATATCTCCATGAACAGGGGATTCCTGTGCTCTCTGTGGGATTCATTCAGGGGTCCAG 366
 QY 171 PheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArgGluLeuGluThrHis 190
 Db 367 TTTGTGGCTTCTGAGCCCACTGGATTTTATCTATATTCAGAGAGCTGGAACGGAT 426
 QY 191 GlySerAsnLeuThrGluGluGlnLeuGluThrHisThrIleSerAlaThrIysGluAla 210
 Db 427 GGCTCAAACTTCACAGAGGAACAGCTTGAACACACACTATATCTGCGTGGAAAGAGCT 486
 QY 211 LysArgGlnThrAsnGlyArgAsnAspSerGlnTyrArgProGlnGlnHisLeuValHis 230
 Db 487 AAGCGGCAAACTTATGGAAGAAATGATGSAACAACCTAGATCAATCAGCATCTAGTGCT 546
 QY 231 AlaThrProTyrGluSerLeuArgAspIleAlaValIysLeuLeuGlnAsnGlyIleSer 250
 Db 547 GCCACCCCTTATGAATCTTGGAGGGGTATTGCCATGAATAACTTGAACCTGGTATTCT 606
 QY 251 ThrValProValIleTyrSerSerSerSerAspGlySerPheProGlnLeuLeuHisLeu 270
 Db 607 ACAGTACCAATCATCTATTATCGTCATCAGATGATGATGCTTCCGACGCTGTGATCTT 666
 QY 271 AlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPheLysAsnSerThrGlyAsn 290
 Db 667 GCATCCCTTTTCAGGAATTTTGAATGATACCTGTAGTACTTTAAGACTTCCACTGGTNGT 726
 QY 291 LeuProfile 293
 |||||

Db 727 TTGCCGATT 735

RESULT 7

CD903099
 LOCUS 712 bp mRNA linear EST 14-JUL-2003
 DEFINITION G356.109C03F010918 G356 Triticum aestivum cDNA clone G356109C03,
 mRNA sequence.
 ACCESSION CD903099
 VERSION CD903099.1 GI:32677427
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 712)
 AUTHORS Genoplante,
 TITLE Genoplante, a major partnership french program in plant genomics
 JOURNAL Unpublished (2003)
 COMMENT Contact: Genoplante

Genoplante
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french
 plant genomics programme "Genoplante" (<http://www.genoplante.com>)
 and <http://genoplante-info.infobiogen.fr>.
 Location/Qualifiers
 1..712
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultiivar="recital"
 /db_xref="taxon:4565"
 /clone="G356109C03"
 /tissue type="grain (356 degrees per day after
 pollination)"
 /clone_lib="G356"

FEATURES
 source

Alignment Scores:
 Pred. No.: 1,44e-112 Length: 712
 Score: 1050.00 Matches: 203
 Percent Similarity: 94.30% Conservative: 12
 Best Local Similarity: 89.04% Mismatches: 12
 Query Match: 44.83% Indels: 1
 DB: 14 Gaps: 0

ORIGIN

US-09-857-525C-2 (1-451) x CD903099 (1-712)
 QY 203 ThrIleSerAlaTyrLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTyr 222
 Db 12 ACTATATCTCGGTGGAAGAGGCTAAGCGCAACTTATGGAAGAAATGATGGCAACT 71
 QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
 Db 72 AGATCAATACAGCATCTAGTGCATGCCACCTTATGATCTCTGAGGGGTATGCCATG 131
 QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly 262
 Db 132 AAAATACTCGAAACTGGCATTTCTACAGTCCCAATCATCTATTCATCGATCAGATGGA 191
 QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
 Db 192 TCGTTTCCGAGCTGTGTCATCTTTCATCCCTTTTCAGGAATTTGAAATGATCTCTAGA 251
 QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
 Db 252 TACTTCAAGAACTCCACTGGTAGTTTCCCGATTTCTAAACCAACAGATGATGCTCAATTCG 311
 QY 303 LeuGlySerThrValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
 Db 312 CTGGGTACCTGGGTTCAAAAAATGGTGAAACCAAAATGGTCATCCATTTGGCTATGTTCGG 371

323	Qy	ProAsnAlaSerLeuSerSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle	342
372	Db	CCTAATAACATCTCTTAGCTCTGCCCTTAACCTGTGTGGTTCAAGCTGGAGTTAGTTCAATA	431
343	Qy	ProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla	362
432	Db	CCCATTGTGGATGATAACGACTCGCTGATCCACACATATCTCCAGAAGTGACATCACAGCT	491
363	Qy	LeuAlaLysAspLysValTyrHisValArgLeuAspGluMetThrIleHisGlnAla	382
492	Db	CTAGCGAAGACAAGGCTACACCCCAATATCCGCTAGATGAGATGACCATTCATCAGGCC	551
383	Qy	LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln	402
552	Db	TTGCAGCTCGGGCAAGACGCGAATTCACCTTTTGGACCTTTTCAATGGTCRAAGATGCCAG	611
403	Qy	MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAspProGlyVal	422
612	Db	ATGTGTCTCCAGTCTGACCCCTTTCGTGAAGAGTTATGCAGAGATTGGCTATCTCTGGGGT	671
422	Qy	ArgValPheIleValGlu	429
672	Db	GCCTCGCGTGTTCATCTCGTGA	693

RESULT 8	CA262684	693 bp	linear	EST 26-SEP-2003
LOCUS	CA262684		mrna	
DEFINITION	SCPIB2024A07.g		Saccharum officinarum	cDNA clone SCPIB2024A07
			5', mRNA sequence.	

ACCESSION	CA262584
VERSION	CA262584.1
KEYWORDS	EST.
SOURCE	Saccharum officinarum
ORGANISM	Saccharum officinarum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD; Gramineae; Panicoideae; Andropogoneae; Saccharum.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 683)	Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.	The libraries that made SUCEST	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)	
	Contact: Arruda p			
	Centro de Biologia Molecular e Engenharia Genetica			
	Universidade Estadual de Campinas			
	Caixa Postal 6010, 13083-970, Campinas SP, Brazil			
	Tel: 55 19 3788 1137			
	Fax: 55 19 3788 1089			

Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bccccenter.fcav.unesp.br>
Plate: 024 row: A column: 07
Seq primer: T7 promoter Primer.

```

FEATURES
source
Location/Qualifiers
1. .683
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCPILB2024A07"
/lab_host="DH10B"
/clone_lib="lB2"
/note="Organ: Lateral buds from plants adult plants
growing in greenhouse; Vector: pSport1; Site 1: SalI;
Site 2: NotI; An unidirectional cDNA library generated
from [lateral buds from plants adult plants growing in
greenhouse]. cDNA was prepared from polyA+ mRNA using
SuperScript plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public."

```

ORIGIN

Alignment Scores:		
Pred. No.:	5.18e-112	Length: 683
Score:	1045.00	Matches: 205
Percent Similarity:	94.52%	Conservative: 2
Best Local Similarity:	93.61%	Mismatches: 12
Query Match:	44.62%	Indels: 1
DB:	14	Gaps: 0

US-09-857-525C-2 (1-451} x CA262684 (1-683)

Qy 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheClnAlaIleCysSer 22
 |||||
 Db 26 GACATTTGTCGATGTCCTCTGGT - AGTTGCCCACTGATTTTCAGGCATTTCGACG 84
 |||||

Qy 23 LeuSerProGlyIleHisGluTyrIysPhePheValAspGlyGluTrpArgHisAspGlu 42
Db 85 CTGTCTCCAGGGAATCAGCAGTACACAGCCTTTTGTGACGGGAGTGGCGGCATGATGAG 144

Qy 43 ArgCinProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyLeuThrArgGlu 62
145 CGCCACCTACTATATCCGGGAGTTCGTAGTTAACACATTCTACTTGACAGAGAA 204
Db

Qy 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 :::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Dd 205 TTTAAACCAATTAAGCGCTTATTAAATCCAGCACCATTTGAAGCAGGATGGACATGGAT 264

Qy 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 265 GTGGATTAATGAAATTTTCACGACGGTTACGTTGTCGGATGGTACTGTTTCGGGAGGT 324
 Db

Qy 103 ThrLeuArgValSerGluAlaAlaIleClnIleSerArgCysArgValSerGluTyrLeu 122
 ||| ||| ||| ||| ||| ||| |||
Dd 325 ACTCCGAGAGTTTCAGAGGCTGCATACAATCTCTAGGTGCCGGCTTCTCGAATACTG 384

Qy 123 AsnLeuHisThrCysTyrAspLeuProAspSerGlyLysValIlealaLeuAspIle 142
Db 385 AATTTGCATACATGCTATGATTTACTCCGGATTCCGGCAAGTTATTGCTCTGGACATT 444

[illegible]

183	IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluLeuThrHis	202
QY		
163	LeuThrAspSerPheArgGlyGlnPheValIleLeuLeuSerProLeuAspPheLeuLeu	181
QY		
505	CTGGGACATTCACAGGTCAATTTGTTGGTCTTCTGAGCCCATTCGATTTCACTT	564
Db		

203	ThrIleSerAlaTrpLysGluAlaLysAroGlnThrAsnGlyArgAsnAspSerGln	221
565	ATATTGGCGAGCTAGAACTCATGGCTCGAACCTGCAGAGAAGCAGCCTTGAACACAC	624
Db		
203	ThrIleSerAlaTrpLysGluAlaLysAroGlnThrAsnGlyArgAsnAspSerGln	221
565	ATATTGGCGAGCTAGAACTCATGGCTCGAACCTGCAGAGAAGCAGCCTTGAACACAC	624
Db		

Db 625 ACTATATCTCATCGAAGAGGCTAAGCGGCACCTTATTCGAAGAAATGATGTCAG 681

LOCUS	BU098591	601 bp	mrna	linear	EST 29-AUG-2003
DEFINITION	946136Fl.y1 946 - tassell primordium prepared by Schmidt lab				
	Zeas				
	mays cDNA. mrna sequence.				

maple, cedar, maple sequence.

ACCESSION BU098591
VERSION BU098591.1 GI:22546280
KEYWORDS EST

KEYWORDS	SOURCE	ORGANISM
ZSL.	Zea mays	Zea mays

REFERENCE
1 (bases 1 to 601)

AUTHORS	Walbot, V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL	University Unpublished (1999)

COMMENT

Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946136 row: F column: 11.
Location/Qualifiers

FEATURES
source

1..601
/organism="Zea mays"
/mol_type="mRNA"
/cultiivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/disage="just after the transition from vegetative to inflorescence development"
/lab_host="XLOLR"
/clone_lib="946 - tassel primordium prepared by Schmidt lab"
/note="Organ: tassels; Vector: HybridZAP; Site 1: EcoRI; Site 2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybridZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

ORIGIN

Alignment Scores:
Pred. No.: 6.25e-109 Length: 601
Score: 1018.00 Matches: 193
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 43.47% Indels: 0
DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x BU098591 (1-601)

Qy 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
Db 21 GAAACATTCGCGAGTCTCTCTATCGAGGCTGCCCATCTGATTTTCAGGCTATTGCGAT 80
Qy 23 LeuSerProGlyIleHisGluCysPhePheValAspGlyGluTrpArgHisAspGlu 42
Db 81 CTGCTCTCTGGGATTCACGAGTACAAATCTTTGTGGACGGGAATGGCGGCATGATGAG 140
Qy 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrlleuThrArgGlu 52
Db 141 CGTCAACCTACCATATCTGGGGAGTTGGCATAGTTAAACACACTTTACTTGACAGGAA 200
Qy 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
Db 201 TATAACCAATAAATACACTTATCAAGTCCAGCACCTGGAGCAGGATGAACATGGAT 260
Qy 83 ValAspAsnGluAsnPheGluArgThrValThrLeuSerAspGlyThrValSerGluGly 102
Db 261 GTGGATAATGAAATTTTCAAGTACGGTTACGTTGTGATGATGGTACCGTTTCAGAAAGT 320
Qy 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrlleu 122
Db 321 ACTCTGAGAGTTTCAGAGCTCAATACAAATATCTAGGTGCGTGTCTGAATATCTG 380
Qy 123 AsnLeuHisThrCysTyAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
Db 381 AATTTGATACATGCTATGATTTACTCCAGATTTCTGGCAAGTTATTGCCCTAGACATT 440
Qy 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
Db 441 AATTTTACCTGTGAGCAATCATTTCCATATTTCTCCATGAACAGGGGATTCCTGTAGCTCT 500
Qy 163 LeuTrpAspSerPheArgGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db 501 CTCTGGGACTCATTCAGAGGTCAATTTGTTGTTCTCTTAGCCCATTTGGATTTCTATCTC 560

Qy 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThr 195
Db 561 AATTCGGGAGCTAGAAACTCATGGCTCGAATTGACA 599

RESULT 10

BE598304

LOCUS

DEFINITION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Cordnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions

below Phred quality 16. The threshold for highest quality sequence

is 20.

Seq primer: JEN REV

High quality sequence stop: 546

POLYA=No.

Location/Qualifiers

1..595

/organism="Sorghum bicolor"

/mol_type="mRNA"

/db_xref="taxon:4558"

/clone_lib="Pathogen induced 1 (PII)"

/note="Organ: Anthracnose-infected leaves from

two-week-old sorghum plants 48 hr after inoculation;

Vector: pBluescript II from Lambda Zap II; Site 1: XhoI;

Site 2: EcoRI; Two-week-old sorghum plants (BTK 623

cultivar) were infected with pathogen (isolate PRM421 of

Colletotrichum graminicola, which is a sorghum isolate).

RNA was prepared from infected leaves harvested from 45

seedlings 48 hours after inoculation. Note: young

seedlings (2 weeks old) exhibit juvenile resistant

reaction, which is an incompatible interaction. As they

grow older (4 weeks or older), plants resume susceptibility

to anthracnose disease. The library was made from poly-A

RNA in the cloning vector lambda Zap II. Clones to be

sequenced were prepared by mass excision. WARNING: While

most or all ESTs are expected to derive from the host

plant, no effort was made to eliminate ESTs deriving from

the pathogen."

the pathogen."

ORIGIN

Alignment Scores:

Pred. No.: 1.07e-106 Length: 595

Score: 999.00 Matches: 190

Percent Similarity: 96.95% Conservative: 1

Best Local Similarity: 96.45% Mismatches: 6

Query Match: 42.66% Indels: 0

DB: 10 Gaps: 0

US-09-857-525C-2 (1-451) x BE598304 (1-595)

Qy 119 SerGluTyrlleuAsnLeuHisThrCysTyAspLeuLeuProAspSerGlyLysValIle 138
Db 3 ACGAGGTATCAATTTGCATACATGCTATGATTTACTCCCCGGATTCGGGCAAGGTATT 62

```

Qy 139 AlaLeuAspIleAsnLeuProVallyLeuGlnSerPheHisIleLeuHisGluGlnGlyLeu 158
Db 63 GCTCTGGACATTAATTTACCTGTGAAGCAATCTTTCATATTCCATGACAGGGGATT 122
Qy 159 ProValAlaProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeu 178
Db 123 CCTGTAGTCTCTCTGGAGACTCATTTAGAGGTCAATTTGTGTCTCTCTGAGCCATTG 182
Qy 179 AspPheIleLeuLeuLeuGluLeuGluThrHisGlySerAsnLeuThrGluGln 198
Db 183 GATTTTCATATTATTTGGGGAGCTAGAAACTCATGGCTCGAACTTGACAGAGAGACAG 242
Qy 199 LeuGluThrHisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsn 218
Db 243 CTGGAACACATACGATATCTGCATGGAAGAGGCTAAGCGGCAAACTTATGAGAGAAT 302
Qy 219 AspSerGlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGlnSerLeuArg 238
Db 303 GATGGTCAATGGCGCACCATCAGCATCTAGTGCATGCTACCCCTTATGAGTCTTTGAGG 362
Qy 239 AspIleAlaVallyLeuLeuGlnAsnGlyLeuSerThrValProValIleTyrSerSer 258
Db 363 GACATTCGAGTAAACTTTTGCAAAATGGGCAATTTCTACAGTGCCTGATTTATTATCATCA 422
Qy 259 SerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyLeuLeuLys 278
Db 423 TCATCGGATGGATCATTTCCCGCAGTTGTTGCATCTTGCATCCCTCTCTGGAAATTTTGAAA 482
Qy 279 CysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProLeuLeuAsnGlnProVal 298
Db 483 TGTATTTGTAGATATTTCAAAACCTCGACTTGTAATTTGCCTATTCTGAAACCAACACAGT 542
Qy 299 CysSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSer 315
Db 543 TGCTCATTTCCACTGGGTTCTGGGTTCGAAATTTGGTATCTTCACTACAGT 593

```

```

RESULT 11
LOCUS CA218557 577 bp mRNA linear EST 25-SEP-2003
DEFINITION SCUFADIC09B02.g Ad1 Saccharum officinarum cDNA clone SCUFADIC09B02
ACCESSION CA218557
VERSION CA218557.1 GI:35269977
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum.
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: C09 row: B column: 02
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .577
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCUFADIC09B02"
/lab_host="DH10B"

```

FEATURES

```

source BE123279
DEFINITION BE123279
LOCUS 945040D06.Y1 945 - Mixed adult tissues from Walbot lab, same as 797
(SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE123279
VERSION BE123279.1 GI:8516554
KEYWORDS EST.
SOURCE Zea mays

```

```

/clone_lib="AD1"
/note="Organ: seedlings inoculated with Gluconacetobacter
diazotrophicans; Vector: pSport1; Site_1: SalI; Site_2:
NotI; An unidirectional cDNA library generated from
[seedlings inoculated with Gluconacetobacter
diazotrophicans]. cDNA was prepared from polyA+ mRNA using
SuperScript plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3.49e-103 Length: 577
Score: 969.00 Matches: 184
Percent Similarity: 97.38% Conservative: 2
Best Local Similarity: 96.34% Mismatches: 5
Query Match: 41.37% Indels: 0
DB: 14 Gaps: 0

US-09-857-525C-2 (1-451) x CA218557 (1-577)

Qy 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerLeu 23
Db 1 CATTTGCCGATCTCTCTGTGGAAGGCTGCCCACTGATTTTCAGGCTATTTGCAGCTG 60
Qy 24 SerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrPheHisAspGluArg 43
Db 61 TCTCCAGGGATCCACGAGTACAGTCTTTTGTGGACGGGAGTGGCGCATGATGAGCGC 120
Qy 44 GlnProThrIleSerGlyGluPheGlyLeuValAsnThrLeuTyrLeuThrArgGluTyr 53
Db 121 CACCTTACTATATCCGGGGAGTTTGGTATAGTTAAACACATTTGTACTTTGACAAGAGAAATTT 180
Qy 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83
Db 181 AACCAATAAACGCTTATTAAAGTCCAGCACAAATTCGAAGCAGATGACATGATGTG 240
Qy 84 AspAsnGlnAsnGlnGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
Db 241 GATAATGAAATTTTCAACGCTACGTTACGTTGTGGATGGTACTGTTCGGAAGGTACT 300
Qy 104 LeuArgValSerGlnAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsn 123
Db 301 CCAGAGATTTCCAGAGGCTGCATACAAATCTTAGGTGCCCGCTTCTGAATATCTGAAT 360
Qy 124 LeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
Db 361 TTGCATACATGCTATGATTTTACTCCCGGATTCAGGCAAGGTTATTTGCTTGGACATTAAT 420
Qy 144 LeuProVallyGlnSerPheHisIleLeuHisGluGlnGlyLeuProValAlaProLeu 163
Db 421 TTACCTGTGAAGCAATCTTTTCATATTTCTACATGACACAGGGGATTCCTGTAGCTCCTCTG 480
Qy 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183
Db 481 TGGACATCATTCAGAGTCAATTTGTGTGTTCTTAGGCCCATTTGGATTTCTACTATTATA 540
Db 541 TTGGGGAGCTAGAAACTCATGTGCTTGAACCTG 573

```

RESULT 12

```

BE123279
LOCUS 945040D06.Y1 945 - Mixed adult tissues from Walbot lab, same as 797
DEFINITION (SK) Zea mays cDNA, mRNA sequence.
ACCESSION BE123279
VERSION BE123279.1 GI:8516554
KEYWORDS EST.
SOURCE Zea mays

```


ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
C-ade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 559)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 945040 row: D column: 36.
FEATURES
Location/Qualifiers
Source 1..559
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/tissue_type="tassel, kernal, silk, husk, root, leaf"
/dev_stage="fully-grown"
/lab_host="DH10B"
/clone_lib="945 - Mixed adult tissues from Walbot lab,
same as 707 (SK)"
/note="Organ: tassel, kernal, silk, husk, root, leaf;
Vector: pGAD10; Site 1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator plant.
Tissue ratio is 4:2:1:1:1 (tassel, kernel, silk, husk,
root, leaf). Unidirectionally cloned. New library number
given to library 707 for additional sequencing."

ORIGIN
Alignment Scores:
Pred. No.: 1.28e-102 Length: 559
Score: 964.00 Matches: 186
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.16% Indels: 0
DB: 10 Gaps: 0

US-09-857-525C-2 (1-451) x BE123279 (1-559)

Qy 66 IleAsnThrLeuSerProSerThrProGlySerArgMetAsnMetAspValAspAsn 85
Db 2 ATAAACACCTTATCAAGTCCAGCACACCTGGAGCAGGATGAACATGGATGATAAT 61
Qy 86 GluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArg 105
Db 62 GAAATTTTCAAGTACGGTTACGTTGTCAGATGGTACCGTTTCAGAGGACTCTGAGA 121
Qy 106 ValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHis 125
Db 122 GTTTCAGAGCGCTGCAATACAAATATCTAGGTCGTCGTTCTGAAATCTCAATTTGCAT 181
Qy 126 ThrCysTyrAspLeuLeuProAspSerGlyValIleAlaLeuAspIleAsnLeuPro 145
Db 182 ACATGCTATGATTTTACTCCCAAGATTTGGCAAGGTTATTCGCCCTAGACATTAATTTACCT 241
Qy 146 ValLysGlnSerPheHisIleLeuHisGlnGlnGlyIleProValAlaProLeuTyrAsp 165
Db 242 GTGAAGCAATCTTCATATTTCTCATGAAACAGGGGATTCCTGTAGCTCTCTCTGGAC 301
Qy 166 SerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArg 185
Db 302 TCATTCAGAGGTCAATTTGTTGGTCTCTCTAGCCCATTTGGATTTTCATATCATATTCGG 361
Qy 186 GluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThrIleSer 205
Db 362 GAGCTAAGAACTCATGGCTCGAATCTTGACAGAGAGCAGCTTGAACACACACTATATCT 421

Qy 206 AlaTyrLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTyrArgProGln 225
Db 422 GCATGGAAGAGGCTAAGCGGCAACAACTAATGAGAAATGATAGTCAGTGGCGCCGCA 481
Qy 226 GlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeuLeu 245
Db 482 CAGCATCTAGTGCATGCCACCCCTTATGAGTCTTGGAGGACATTCGAGTAAGCTTTTG 541
Qy 246 GlnAsnGlyIleSerThr 251
Db 542 CAAATGGCATTTCTTACA 559
RESULT 13
LOCUS BU499410
DEFINITION 946174G09.yl 946 - tassel primordium prepared by Schmidt lab Zea
mays cDNA, mRNA sequence.
ACCESSION BU499410
VERSION BU499410.1 GI:22819320
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 549)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946174 row: G column: 09.
FEATURES
Location/Qualifiers
Source 1..549
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XL0LR"
/clone_lib="946 - tassel primordium prepared by Schmidt
lab"
/note="Organ: tassels; Vector: HybridZAP; Site 1: EcoRI;
Site 2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybridZAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."

ORIGIN
Alignment Scores:
Pred. No.: 1.63e-102 Length: 549
Score: 963.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.12% Indels: 0
DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x BU499410 (1-549)

Qy 10 IleGluGlyCysProThrValPheGlnAlaIleCysSerLeuSerProGlyIleHisGlu 23
Db 1 ATCGAAGGCTGCCACCTGATTTTCAGCTATTTTCAGCTATTTTCAGCTATTTTCAGCT 60
Qy 30 TyrLysPhePheValAspGlyGluTyrArgHisAspGluArgGlnProThrIleSerGly 49

Db 61 TACAAATCTTTGTGGACGGGATGCGGATGATGAGCTCAACCTACCATATCTGGG 120
 Qy 50 GlupheglyleValAsnThrLeuTyrLeuThrArgGluTyrAsnGlnIleAsnThrLeu 69
 Db 121 GAGTTTGGCATAGTTAAACACACTTTACTTGACAAAGGGAATATAACCAATAACACCTTA 180
 Qy 70 SerSerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAsnPheGln 89
 Db 181 TCAAGTCCAAAGCACACCTCGAAGCAGGATGAGCAATGATGGATGAATGAATAATTTTCAA 240
 Qy 90 ArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGluAla 109
 Db 241 CTTACGGTTACGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 Qy 110 AlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTyrAsp 129
 Db 301 GCAATACAAATATCTAGGTGTGCTGTGTTCTGAAATATCTGAAATATGCAATATGCTATGAT 360
 Qy 130 LeuLeuProAspSerGlyLeuValIleAlaLeuAspIleAsnLeuProValIleGlnSer 149
 Db 361 TTACTCCAGATCTTGGCAGAGGTATGTCCTTACAGATGATGATGATGATGATGATGATGAT 420
 Qy 150 PheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTyrAspSerPheArgGly 169
 Db 421 TTCCATATCTCCATGACACAGGGGATCTCTAGCTCTCTCTGGACTCAATCAGAGST 480
 Qy 170 GlnPheValGlyLeuLeuSerProLeuAspPheIleIleuLeuLeuArgGluLeuThr 189
 Db 481 CAATTTTGTGGTCTCTCTAGCCCATGGAATTTCTACTATATTTGGGGAGCTAGAACT 540
 Qy 190 HisGlySer 192
 Db 541 CATGGCTCG 549

RESULT 14
 CA253676
 LOCUS
 DEFINITION SCRLFL4103H07.g FL4 Saccharum officinarum cDNA clone SCRLFL4103H07
 5', mRNA sequence.
 ACCESSION
 VERSION CA253676
 KEYWORDS EST.
 SOURCE
 ORGANISM
 Saccharum officinarum
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum.
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 The libraries that made SUCEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccenter.fcav.unesp.br
 Plate: 103 row: H column: 07
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers
 1. 631
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCRLFL4103H07"
 /lab_host="DH10B"
 /clone_lib="FL4"
 /note="Organ: Developed inflorescence and rachis

FEATURES

source
 BE704511
 DEFINITION SC01.08c11 A SC01 AAFC ECORC cold stressed winter rye seedlings
 Secale cereale cDNA clone SC01.08c11, mRNA sequence.
 ACCSSION BE704511
 VERSION BE704511.1 GI:10092881
 KEYWORDS EST.
 SOURCE Secale cereale (rye)
 ORGANISM Secale cereale
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

(20cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI;
 An unidirectional cDNA library generated from [Developed
 inflorescence and rachis (20cm-long)]. cDNA was prepared
 from polyA+ mRNA using Superscript Plasmid System Kit
 (Invitrogen). The double-strand cDNAs were fractionated
 in a sepharose CL-2B 40cm-columns and fragments sizing
 between 0.8 and 1.5 kb were directionally cloned into the
 vector. Details of each source of RNA and library
 construction can be obtained at
 http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:
 Pred. No.: 3,18e-101 Length: 631
 Score: 953.00 Matches: 187
 Percent Similarity: 98.45% Conservative: 4
 Best Local Similarity: 96.39% Mismatches: 3
 Query Match: 40.69% Indels: 0
 DB: 14 Gaps: 0
 US-09-857-525C-2 (1-451) x CA253676 (1-631)
 Qy 258 SerSerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyLeuLeu 277
 Db 2 TCATCATCAGATGGATCATTCCCGTTTTTGTTCATCTTGCATCCCTTCTCGAATGTG 61
 Qy 278 LysCysIleCysArgTyrPheIysAsnSerThrGlyAsnLeuProIleuAsnGlnPro 297
 Db 62 AAATGTATTGTGATATATTCAAAACACCGACTGGTAATTTGCCCTATTCTGAACCAACA 121
 Qy 298 ValCysSerIleProLeuGlySerTyrValProIleGlyAspLeuAsnSerArgPro 317
 Db 122 GTGTGCTCCATTCGCTGGGTTCCTGGTTCCGAAATTTGGTATCTTAACAGTCTCCA 181
 Qy 318 LeuAlaMetLeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAla 337
 Db 182 TTGGTATGTTGGACCTTAATGTCATCACTAGCTCTGCCCTTAACATGTTGGTCAAGCT 241
 Qy 338 GlyValSerSerIleProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArg 357
 Db 242 GGAGTAAAGCTCAATACCAATCTGGATGAAATACGCCCTGCTTGACACTTACTCTAGA 301
 Qy 358 SerAspIleThrAlaLeuAlaIysAspLysValTyrThrHisValArgLeuAspGluMet 377
 Db 302 AGTGACATCACAGCCCTAGCAGAGACAGCTCTACACATGTTCCGCTGGATCAGATG 361
 Qy 378 ThrIleHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsn 397
 Db 362 ACCATTTCATCAGGCTTTGCAGCTTGGACAGATGCCAATACACCTTTTGGATTTTAAAC 421
 Qy 398 GlyGlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeu 417
 Db 422 GGTGAGAGATGCCAGATGTCCTCCGATCTCATCTTTGCGAAGGATGATGAGGACATG 481
 Qy 418 AlaAsnProGlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGly 437
 Db 482 GCTAATCTGNGGTGCGCGGGGTGTTCTATTGGAAGCTGGAGCAAAACGTTGGAGGCG 541
 Qy 438 IleIleSerLeuSerAspIlePheIysPheLeuLeuSerLeu 451
 Db 542 ATTATATCACTAGTATATTTTCAAGTTCTTGTGTAACCTG 583

RESULT 15

BE704511/c
 LOCUS
 DEFINITION SC01.08c11 A SC01 AAFC ECORC cold stressed winter rye seedlings
 Secale cereale cDNA clone SC01.08c11, mRNA sequence.
 ACCSSION BE704511
 VERSION BE704511.1 GI:10092881
 KEYWORDS EST.
 SOURCE Secale cereale (rye)
 ORGANISM Secale cereale
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

US-10-425-114-1948

Alignment Scores:

Pred. No.: 3.55e-268 Length: 1844
 Score: 2231.00 Matches: 429
 Percent Similarity: 97.33% Conservativity: 8
 Best Local Similarity: 95.55% Mismatches: 12
 Query Match: 95.26% Indels: 0
 DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-425-114-1948 (1-1844)

QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 Db 206 GAACATTGGCGATGCTCTCGTGAAGGTCGCCACTGTATTTCAGGCTATTTCAGC 265
 QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu 42
 Db 266 CTGTCTCCAGGATTCATGAGTACAGTTCTATGTGACCGGGAGTGGCGTCATGATGAG 325
 QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
 Db 326 CGCCAACTTACTATATCTGGGAGTTCGGTATAGTTAAACACACTATATTCGACAGGGAA 385
 QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 Db 386 TTTAATCTAAATAATGCTTTATTAAATCCAAAGCACACCTGGAAGCAGGATGAACATGGAT 445
 QY 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 Db 446 GTGGATAATGMAACTTTCAACATACGTTACATTTGCGATGTGATCCATTCAGAAAGGT 505
 QY 103 ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
 Db 506 ACTGTGAGAGTTTCAGAGGCTGCAATACAAATCTCAAGGTGCGCGTTCGTGATATCTC 565
 QY 123 AsnLeuHisCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
 Db 566 AATTTCATACATGCTATGATTTACTCCCGGATTCGGCAAGGTTATTGCTCTGACATT 625
 QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 Db 626 AATTTCATGCAACATCTTTTCAATCTGTCATGAACAGGGATTCCTGTAGTCTCT 685
 QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
 Db 686 CTCTGGGACTATTTCAGAGGTCATTTGCGCCCTTCGAGCCCATGGAATTCATCT 745
 QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
 Db 746 ATATTGCGGGAGCTAGAAACTCATGGCTCGAACCTTGACAGAGATCAGCTTGAACACAC 805
 QY 203 ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
 Db 806 ACTATATCTGCATCGAAAGAGGCTAAGCGCCAACTTGTGAAGAAATGATGTCAGTGG 865
 QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
 Db 866 CGACACATCAGCATCTAGTCATGTCACCCCTTATGAGTCCTTCAGGACATTCGACGTA 925
 QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAspGly 262
 Db 926 AAGCTTTTGCTAAATGACATTTTCACAGTGCAGTTATTTATTCATCATCATCATGATGA 985
 QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
 Db 986 TCATTCCCTCAGTTATTGACCTTGCATCCCTTCTGGAATTTGAAATGATTTATTAGG 1045
 QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
 Db 1046 TATTTTAAAAAATCAACTGTTAATTTGCTTATTTCTGAACCAACCGGTGTGCTCATTCG 1195
 QY 303 LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322

Db 1106 CTGGTTCTCTGGTTCCGAAATCGTGATCCAAACAGTGTGCTCCATTGGCTATGTTGGA 1165
 QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
 Db 1166 CTTAATGCAUCATTAGTCTGCGCTTAACATGTTGGTTCGAAGCTGGAGTGAGCTCAATA 1225
 QY 343 ProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
 Db 1226 CCAATTGTGGATGAAACGACTCCCTGCTTGACACTTACTCTAGAGTGAATCATCACAGCC 1285
 QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
 Db 1286 CTAGCTAAAGCAAGGCTTACACACATGTTCCGCTGGATGAGTACCACTTCATCAGGCT 1345
 QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
 Db 1346 TTACAGCTTGGACAGATGCCATACGCTTTTGGATTTTTTAACGGTTCAGAGATGCCAG 1405
 QY 403 MetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyVal 422
 Db 1406 ATGTGCTCCGGTCTGATCCTTTGCTGAAGGTGATGAGCGACTGGCTAATCCGGGGTG 1465
 QY 423 ArgArgValPheIleValGlnAlaGlySerIysArgValGluGlyIleIleSerLeuSer 442
 Db 1466 CGCGGGTGTTCATAGTAGAAGCTGGGAGCAACAGTGTGAGGGTATTATATCATCTGAGT 1525
 QY 443 AspIlePheLysPheLeuLeuSerLeu 451
 Db 1526 GATATTTCAAGTTCTTGCTGAGCTTG 1552

RESULT 2

US-10-425-114-27250

; Sequence 27250, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E.

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 27250

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB4729-018-A12_FLI

US-10-425-114-27250

Alignment Scores:

Pred. No.: 6.69e-268 Length: 2228
 Score: 2230.00 Matches: 428
 Percent Similarity: 97.33% Conservativity: 9
 Best Local Similarity: 95.32% Mismatches: 12
 Query Match: 95.22% Indels: 0
 DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-425-114-27250 (1-2228)

QY 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
 Db 589 GAACATTGGCGATGCTCTCGTGAAGGTCGCCACTGTATTTCAGGCTATTTCAGC 648
 QY 23 LeuSerProGlyIleHisGluTyrLysPhePheValAspGlyGluTyrArgHisAspGlu 42
 Db 649 CTGTCTCCAGGATTCATGATGATCAAGTTCATGTGACCGGGAGTGGCGTTCATGATGAG 708

QY 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGlu 62
 Db 709 CGCAACCTACTATATCTGGGAGTTCGGTATAGTTAAACACACTATATCTTGACAGGGAA 768
 QY 63 TyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAsp 82
 Db 769 TTTAATCAATTAATGAGCTTATTAATCAAGCAACCTGGAAGCAGGATGAACATGGAT 828
 QY 83 ValAspAsnGluAsnThrGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
 Db 829 GTGGATAATGAAATTTCAACATACCGTTACATGTCGATGATGATGATGATGATGATGAT 888
 QY 103 ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
 Db 889 ACTGTGAGAGTTTCAGAGGCTGCAATACAAATCTCAAGGTGCGCGTTCTGTAGTATCTC 948
 QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyValIleAlaLeuAspIle 142
 Db 949 AATTTCATACATGCTATGATTTACTCCCGATTCGGGCAAGGTTATGTCTCTGGACATT 1008
 QY 143 AsnLeuProValIysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaPro 162
 Db 1009 AATTACCTGTGAAGCAATCTTTCAATTCATGTCATGCAAGGGATTCCTGTAGCTCT 1068
 QY 163 LeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
 Db 1069 CTCGTGGATCTCATTCAGAGGTCAATTTGTGGCCTTCAGAGCCATTGGATTTCACTTT 1128
 QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
 Db 1129 ATATTGGGGAGCTAGAACTCATGGCTCGACTTCAGAGATCAGCTTGAACACAC 1188
 QY 203 ThrIleSerAlaTrpIysGluAlaIysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
 Db 1189 ACTATATCTGCATGAAAGAGGCTAAGCGGCAAACTTGTGGAAGAATGATGTCAGTGG 1248
 QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
 Db 1249 CGAGCAGATCAGCATCTAGTGTGATGTCACCCCTTATGAGTCTTGAGGACATTGCACTA 1308
 QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerGly 262
 Db 1309 AAGCTTTTGCTAAATGACATTTCTACAGTGGCAGGTTATTTATCATCATCATCATGSA 1368
 QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuIysCysIleCysArg 282
 Db 1369 TCATTCCCTCAGTTATTCACCTTCATCCCTTCTGGAAATTTGAAATGATTTTATGG 1428
 QY 283 TyrPheIysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
 Db 1429 TATTTTAAACTCACTGGTAAATTTGCTTATTTCTGAACCAACCGGTGTCTCCATTCCG 1488
 QY 303 LeuGlySerTrpValProIysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
 Db 1489 CTGGTTCTCTGGTTCCGAAATCGGTGATCCAAACAGTCGTCTCCATTGGCTATGTGCGA 1548
 QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
 Db 1549 CCTAATGCATCACTTAGCTCTGCTTAAACATGTTGCTCAAGCTGGATGAGTGCATCATA 1608
 QY 343 ProfileValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
 Db 1609 CCAATGTTGATGAAAGAGCTCCCTGCTGACACTTACTCTAGAAAGTGACATCACAGCC 1668
 QY 363 LeuAlaIysAspIysValTyrThrHisValArgLeuAspGluMetThrIleHisGluAla 382
 Db 1669 CTAGCTAAAGACAGGCTACACATGTTGCTCGATGATGATGATGATGATGATGATGATGAT 1728
 QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
 Db 1729 TTAGCTTGGACAAAGATGCAATACGCTTTTGGATTTTAACTGATGATGATGATGATGATGAT 1788

QY 403 MetCysLeuArgSerAspProLeuLeuIysValMetGluArgLeuAlaAsnProGlyVal 422
 Db 1789 ATGTCCTCCGGTCTGATCTCTTCTGGAAGGTGATGAGCGACTGGCTAATCCGGGGGTG 1848
 QY 423 ArgArgValPheIleValGluAlaGlySerIysArgValGluGlyIleIleSerLeuSer 442
 Db 1849 CGCGGGGTGTTCTGATAGAGCTGGGAGCAACGCTGGAGGGTATTATATCACTGAGT 1908
 QY 443 AspIlePheLysPheLeuLeuSerLeu 451
 Db 1909 GATATTTTCAAGTTCTTCTGCTGAGCTTG 1935

RESULT 3
 US-10-425-114-3817
 ; Sequence 3817, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 3817
 ; LENGTH: 1474
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700342215_FLI
 ; US-10-425-114-3817

Alignment Scores:
 Pred. No.: 1,08e-261 Length: 1474
 Score: 2178.00 Matches: 421
 Percent Similarity: 99.76% Conservative: 1
 Best Local Similarity: 99.53% Mismatches: 1
 Query Match: 93.00% Indels: 0
 DB: 13 Gaps: 0

US-09-857-525c-2 (1-451) x US-10-425-114-3817 (1-1474)

QY 29 GluTyrLysPhePheValAspGlyGluTrpArgHisAspGluArgGlnProThrIleSer 48
 Db 2 GAGTACAAATCTTTGTGGACGGGGAATGGCGGCAATGATGAGCGTCAACCTACCATATCT 61
 QY 49 GlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGluTyrAsnGlnIleAsnThr 68
 Db 62 GGGGAGTTTGGCATAGTTAAACACACTTTACTTGACAGGGAATATAACCAATAAACACC 121
 QY 69 LeuSerSerProSerThrProGlySerArgMetAsnMetAspValAspAsnGluAsnPhe 88
 Db 122 TTATCAGTCCAGCACACCTGGAGCAGGATGACATGATGATGATGATGATGATGATGATGAT 181
 QY 89 GlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgValSerGlu 108
 Db 182 CAACGTACGGTTACGTTGTCAGATGTTACCGTTTCAGAAAGTACTCTGAGAGTTTCAGAG 241
 QY 109 AlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsnLeuHisThrCysTyr 128
 Db 242 GCTCAATAACAATATCTAGTGTGCTGTTCTCAATATCTGAGTTGTCATACATGCTAT 301
 QY 129 AspLeuLeuProAspSerGlyIysValIleAlaLeuAspIleAsnLeuProValIysGln 148
 Db 302 GATTTACTCCAGATCTGGCAAGGTTATTGCCCTAGACATTAATTTACCTGTGAGCAA 361
 QY 149 SerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTrpAspSerPheArg 168


```

Db 362 TCATCCATATCTCCATGAACAGGGATCTCTGTAGCTCCTCTCTGGGACTCATTGAGA 421
Qy 169 GlyGlnPheValGlyLeuSerProLeuAspPheLeuLeuLeuLeuLeuLeuLeu 188
Db 422 GGTCATTTGTTGGTCTCTTAGCCCAATGGATTTCACTCATATTTGGGAGCTAGAA 481
Qy 189 ThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThrIleSerAlaTrpIys 208
Db 482 ACTCATGGCTCGAATTTGACAGAGAGACGCTTGAACACACACATATCTGCATGAAA 541
Qy 209 GluAlaIysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGlnHisLeu 228
Db 542 GAGCCTAAGCGGCAAACTAATGAAGAAATCATGCTCAGTGGGACCGCAACAGCATCTA 601
Qy 229 ValHisAlaThrProTyrGluSerLeuArgAspIleAlaValIysLeuLeuGlnAsnGly 248
Db 602 GTGCATGCCACCCCTTATGAGTCTTGAAGGACCTTGCAGTAAAGCTTTTGCAAAATGGC 661
Qy 249 IleSerThrValProValLeuTyrSerSerSerSerAspGlySerPheProGlnLeuLeu 268
Db 662 ATTCTACAGTGCAGTATTTATTTCATCATCATCATGATGATCATTCCTGGGTTCCG 721
Qy 269 HisLeuAlaSerLeuSerGlyIleLeuIysCysIleCysArgTyrPheIysAsnSerThr 288
Db 722 CATCTGTCATCACTTCTCGAATTTTGAATGTATTTGTAGATACTTCAAAACTCAACT 781
Qy 289 GlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpValPro 308
Db 782 GGTAAATTTGCCTATTCTGAACCAACAGTGTCTCTCAATCCCTGGGTTCCG 841
Qy 309 LysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSer 328
Db 842 AAAATGTGTGATCTGAACAGTGTCTCATGTTGCGACCTTAATCATCACTTAGC 901
Qy 329 SerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspAspAsn 348
Db 902 TCTGCCCTTACATGTTGGTTCAGCTGAGTGAAGTCAATACCAATTTGTGGATGACAAC 961
Qy 349 AspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeuAlaIysAspIysVal 368
Db 962 GACTCCCTGCTTGACACTTACTTAGAAGTGAATCACAGCCCTTAGCAAAAGACAGGTC 1021
Qy 369 TyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeuGlyGlnAsp 388
Db 1022 TACACATGTTTGGCTGGAGTGAATGACCATTCACAGGCTTTGCGAGTTGGACAAGAT 1081
Qy 389 AlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMetCysLeuArgSerAsp 408
Db 1082 GCCAATACACCTTTTGGATTTTAAAGCCAGAGATGCCAGATGTGCTCCGCTCTGAT 1141
Qy 409 ProLeuLeuIysValMetGluArgLeuAlaAsnProGlyValArgArgValPheIleVal 428
Db 1142 CCTTGTGAGGTGATGGAGCGACTGGCTAATCTCTGGGTGGCGGGTTCATTGTG 1201
Qy 429 GluAlaGlySerIysArgValGluGlyIleIleSerLeuSerAspIlePheIysPheLeu 448
Db 1202 GAAGCTGGGAGCAACGTGTGGAGGGGATCATATCATCACTAAGTACATTTCAAGTCTTG 1261
Qy 449 LeuSerLeu 451
Db 1262 CTGAGCTTG 1270

```

RESULT 4

```

US-10-437-963-2873
; Sequence 2873, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.

```

```

; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 2873
; LENGTH: 2980
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MBT4530_102608C.1
US-10-437-963-2873

```

```

Alignment Scores:
Pred. No.: 6,99e-232 Length: 2980
Score: 1944.00 Matches: 361
Percent Similarity: 94.50% Conservativity: 34
Best Local Similarity: 86.36% Mismatches: 23
Query Match: 83.01% Indels: 0
DB: 17 Gaps: 0
US-09-857-525C-2 (1-451) x US-10-437-963-2873 (1-2980)

```

```

Qy 3 GluHisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSer 22
Db 685 GAACATTTGCCAATGTCTCTGTGAAGTTGTGCCACCGTTTTTCAGGCCATATGCAGT 744
Qy 23 LeuSerProGlyIleHisGlyIysPhePheValAspGlyGluTrpArgHisAspGlu 42
Db 745 TTGTCCCTCGGATTCATCAGTACAAATTTGTGTGGATGGGAATGGCGGCACGATGAG 804
Qy 43 ArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrIleThrArgGlu 62
Db 805 CGGCACACTCATTAACAGGAGACTATGTTGTGGTGTGTAACTTTATGCTTAGTAGGAC 864
Qy 63 TyrAsnGlnIleAsnThrLeuSerSerProThrProGlySerArgMetAsnMetAsp 82
Db 865 TTGACCCAGATTAATACATATTTGAGCCCTAGTACACCTGGAAAGTAGGATGAACATGGAT 924
Qy 83 ValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGly 102
Db 925 GTGGACACAGCAATTTTCAACGTACTGTTTCTTGTCTGATGGCATTTATTCAGGAAGGT 984
Qy 103 ThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
Db 985 CCTCAGGAATTTTCAGAGGCGAGCTATACAGATCTCTAGGTGTCTGTAGCGGATTTCTG 1044
Qy 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyIysValIleAlaLeuAspIle 142
Db 1045 AATGGCAAACTGGGTATGATTTACTCCAGATTTCTGGCAAGGTTCATGCTCTAGACGTT 1104
Qy 143 AsnLeuProValIysGlnSerPheHisIleLeuHisGlnGlnGlyIleProValAlaPro 162
Db 1105 AATTTCCCTGTGAAGCAATCTTTTCATATTTTCATGTAACAGGGAATTTCTGTGGACCT 1164
Qy 163 LeuTrpAspSerPheArgGlyGlnPheValGlyIleLeuSerProLeuAspPheLeu 182
Db 1165 CTATGGGATTCATTTAGGGGCCAGTTTGTGGCTTTTGGCCCTTGTAGCCCTGATTTTACTC 1224
Qy 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHis 202
Db 1225 ATATTGAGAGAGCTGGAATCTCATGGCTCCATCTGACAGAGAGAGCGCTTGAAACACAT 1284
Qy 203 ThrIleSerAlaTrpIysGluAlaIysArgGlnThrAsnGlyArgAsnAspSerGlnTrp 222
Db 1285 ACTATATCTGGGTGAAGAGGAGGCCAAGCGTCAAACTTATGCCAGAAAACGAGGGTCTTGG 1344
Qy 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
Db 1345 AGGGCAATTCACATTTAGTTTCATGCCACCCCTTATGAATCCCTGAGGGAATTCGTATG 1404

```

```

QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerAspGly 262
Db 1405 AAGATCTGCAAAACGGTGTCTACCGTTCCTCAATATGTTTCTCTCATCCAGATGGC 1464
QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
Db 1465 TCATATCCCAAAATGTTGCATCTTGCTTCCCTTCGGAATTTTGAATGTATTGTAGA 1524
QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
Db 1525 TATTTTAAATAATTTCTCAAGGTAATTTACCTATTTTGGACCACTGTATGCAATTCCT 1584
QY 303 LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
Db 1585 CTGGGTACCTGGGTTCCTCAAAATTTGGTATCTTAATGGCGTCCATTTGGCTATGTTGGG 1644
QY 323 ProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerIle 342
Db 1645 CCTAACACATCTCTTAGCGCTGCCCTGAAATTTGCTGGTTCAAGCTGGTGTGAGCTCAATA 1704
QY 343 ProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
Db 1705 CCAATTTGGATGACAAATGACTCACTGCTTGACACATATTCAGAAAGTGCATACAGCT 1764
QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
Db 1765 CTTGCANAAAGCAAGGTGTACACACATTCGCTTAGATGAGATGACCAATTCATCAGGCT 1824
QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGln 402
Db 1825 CTGAGCTTGGACAGGATGCAAAATTCGCCCTTTGGATCTTTAAACGGCAAGATGTCNG 1884
QY 403 MetCysLeuArgSerAspProLeuLysValMetGluArgLeuAlaAsnPro 420
Db 1885 ATGTGTCTCGGTCTGACACTTTGTTGAAAGTGATGAGCGATTAGCTAATCTCT 1938

RESULT 5
US-10-425-114-33827
: Sequence 33827, Application US/10425114
: Publication No. US20040034888A1
: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovalic, David K.
: APPLICANT: Screen, Steven E.
: APPLICANT: Tabaska, Jack E.
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: CURRENT APPLICATION NUMBER: US/10/425,114
: NUMBER OF SEQ ID NOS: 2003-04-28
: SEQ ID NO 33827
: LENGTH: 1387
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURES:
: OTHER INFORMATION: Clone ID: UC-2MFLM017174G09_FLI
US-10-425-114-33827

Alignment Scores:
Pred. No.: 1,16e-214 Length: 1387
Score: 1802.00 Matches: 352
Percent Similarity: 97.53% Conservative: 4
Best Local Similarity: 96.44% Mismatches: 9
Query Match: 76.94% Indels: 0
DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-425-114-33827 (1-3387)
QY 87 AsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThrLeuArgVal 106

```

```

Db 3 AATTTTCAACATACGGTTACATTGTGGATGGTACCAATCCATCCAGAAGGTACTGTGAGAGTT 62
QY 107 SerGluAlaAlaIleGlnIleSerArgCysArgValSerGluCysIleLeuAsnLeuHisThr 126
Db 63 TCAGAGGCTGCAATACAAATCTCAAGGTGCGGGTTCGTGAGTATCTCAATTTGCAATCA 122
QY 127 CysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsnLeuProVal 146
Db 123 TGTATATGATTTACTCCGGATTTCGGCAAGGTATTATTGCTCTGACATTAATTTACCTGTG 182
QY 147 LysGlnSerPheHisIleLeuHisGlnGlnGlyIleProValAlaProLeuTrpAspSer 166
Db 183 AAGCAATCTTTTCAATTTCTGCATGAACAGGGGATTCCTGTAGTCTCTCTGGACTCA 242
QY 167 PheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIleLeuArgGlu 186
Db 243 TTCAGAGGTCAATTTGTGGCCCTCTGAGCCCATTTGGAATTTCAATCTTATTTGCGGGAG 302
QY 187 LeuGluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHisThrIleSerAla 206
Db 303 CTGAGAACTCTGATGGCTCGAACTTCGACAGAAGATCAGCTTGAACACACACATATATCTGCA 362
QY 207 TrpLysGlnAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgProGlnGln 226
Db 363 TGGAAAGAGGCTAAGCGGCAACTTGTGGAAGAAATGATGTCAGTGCAGCGACACATCAG 422
QY 227 HisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeuLeuGln 246
Db 423 CATCTAGTGCATGCCACCCCTTATGAGTCTTTGAGGACATTCAGATAAGCTTTTGGCAA 482
QY 247 AsnGlyIleSerThrValProValIleTyrSerSerSerAspGlySerPheProGln 266
Db 483 ANTGACATTTCTACAGTGCAGTATTTATTTATTCATCATCATCAGATGATCCTCCTCAG 542
QY 267 LeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyr2helysAsn 286
Db 543 TTATTGACCTTGCATCCCTTTCTGGAAATTTTGAATGTATTTTAGGTATTTTAAATAAC 602
QY 287 SerThrGlyAsnLeuProIleLeuAspGlnProValCysSerIleProLeuGlySerTrp 306
Db 603 TCAACTGGTAATTTGCCATTTCTGAACCAACCGGTGTGCTCCATTCGCTGGGTCTCTGG 562
QY 307 ValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSer 326
Db 663 GTTCCGAAATTCGGTATCCAAACAGTCGTCATTTGGCTATGTTGCGACCTTAATGCATCA 722
QY 327 LeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAsp 346
Db 723 CTTAGCTCTGCCCTTAAATGTTGGTTCAAGCTGGAGTGAGCTCAATATACCAATTTGGAT 782
QY 347 AspAsnAspSerLeuLeuAspThrTyrSerArgSerAspIleThrAlaLeuAlaLysAsp 366
Db 783 GAAACGACTCCCTGCTTGCATCTTACTAGAGTGCATCAGACCCCTAGCTAGTAAGAC 842
QY 367 LysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeuGly 386
Db 843 AAGGTCTACACATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
QY 387 GlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMetCysLeuArg 406
Db 903 CAAGATGCCAATACGCCCTTTTGAATTTTAAACGGTTCAGAGATGCCAGATGTCCTCCGG 962
QY 407 SerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGlyValArgArgValPhe 426
Db 963 TCTGATCCTTTGCTGAAGGTGATGAGGACATGCTGCTATCCGGGGGTGCGCGGGGTGTC 1022
QY 427 IleValGlnAlaGlySerLysArgValGluGlyIleLeuSerLeuSerAspIlePheLys 446
Db 1023 ATAGTAGAAGCTGGGAGCAACGCTGTGAGGGGTATTTATATCACTGAGTGATTTTCAAG 1082
QY 447 PheLeuLeuSerLeu 451

```

Db 1083 TTCTTGCTGAGCTTG 1097

RESULT 6

US-10-424-599-44858

; Sequence 44858, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 44858

; LENGTH: 3077

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT MRT3847_140506C.1

US-10-424-599-44858

Alignment Scores:

Pred. No.:	1-58e-181	Length:	3077
Score:	1542.00	Matches:	306
Percent Similarity:	80.62%	Conservative:	60
Best Local Similarity:	67.40%	Mismatches:	75
Query Match:	65.84%	Indels:	13
DB:	13	Gaps:	4

US-09-857-525C-2 (1-451) x US-10-424-599-44858 (1-3077)

QY 3 GluHisLeuProMetSerProLeuGluGlyCysProThrValPheGlnAlaIleCysSer 22

Db 1019 GAACCTTCTACAAATGTCGCCAGTGAAGGTTGTCCAACTGTTTCAAGTGAATCATAGC 1079

QY 23 LeuSerProGlyIleHisGluTyrlsPheValAspGlyGluTrpArgHisAspGlu 42

Db 1079 TTGCTACCGGTTACCACTACAGTCAAGTATTTGTGTGAGAGATGCGGCGATGCA 1138

QY 43 ArgGlnProThrIleSer-GlyGluPheGlyIleVal-AsnThrLeu-TyrLeuThrArg 61

Db 1139 CATCAACTTATGAGCTGGAGATATGGATAGTCTAACACCTGCTTATTTGGCCACT 1198

QY 62 GluTyrlsGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMet 81

Db 1199 GATCCTAATTACATACCTGTTTAA-CTCCAGACGTTGCTTCTGGAAATAGCATG 1252

QY 82 AspValAspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGlu 101

Db 1253 GATGTGGATATGATGCTTTTCCCGAATGCGCGGTGACCGATGCTACTTTGAGTGAG 1312

QY 102 GlyThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyr 121

Db 1313 GTGCTGCCAAGATATCAGATCTGATGTAACAATATCCGTCAGCGTATTTCTGCATTT 1372

QY 122 LeuAsnLeuHisThrCysTyrlsAspLeuProAspSerGlyIleValIleAlaLeuAsp 141

Db 1373 CTATCTTCACACCCGCTTATGAATTAATCTCCGAGTCAGGCAAGGTTGTGTTGGAT 1432

QY 142 IleAsnLeuProValIleGlnSerPheHisIleLeuHisGluGlnGlyIleProValAla 161

Db 1433 GTTGATCTACCACTGAACACGCAATTTATATTCATATTCATGAGCAGGAGTTTCATGGCT 1492

QY 162 ProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIle 181

Db 1493 CCTCTTTGGGACTCTCTGCAAGGCGCAATTTGTGTGTTCTAGTCTTCGGATTTAT 1552

QY 182 LeuIleLeuArgGluLeuThrHisGlySerAsnThrGluGluGlnLeuGluThr 201

Db 1553 TTAATTTTAAGAGAGCTGGGAATCATGATCCCAATCTGACAGAGAGGACGCTTGAACA 1612

QY 202 HisThrIleSerAlaTrpIleGlyGlnAlaIleCysSer 217

Db 1613 CATACCATATCAGCTTGAAGAGAGGAAATCGTATCTAAATATGACAGCAATGGACAT 1672

QY 218 AsnAspSerGlnTrpArgProGlnGlnHisIleValHisAlaTrpProTyrGluSerLeu 237

Db 1673 GGAACCTGCATTTCAAGA-----TGTTTATCCATCAGGCGCATATGATTAATCTG 1723

QY 238 ArgAspIleAlaValIleLeuGlnAsnGlyIleSerThrValProValIleTyrSer 257

Db 1724 AAGATATTTGCCATGAGATCTTGCAGAGGAGGCTTCAACAGTCTCTATATTCATCA 1783

QY 258 SerSerSerAspGlySerPheProGlnLeuLeuHisIleAlaSerLeuSerGlyIleLeu 277

Db 1784 TCTTCTGAAGATGCTTCATTTCCACAGTTACTACATCTTCTGCTTCACTTTCAGGAATCTT 1843

QY 278 LysCysIleCysArgTyrPheIleAsnSerThrGlyAsnLeuProIleLeuAsnGlnPro 297

Db 1844 AAATGCATTTGTAGTATTTTAGGCACTCTCTAGTTCCTTCTGCTTCACTTCACTTCCA 1903

QY 298 ValCysSerIleProLeuGlySerTrpValProIleGlyAspLeuAsnSerArgPro 317

Db 1904 ATCTGTCAATACCTGTGGGCACTGGGTGCCCAAAATTTGGGAATCAATCGCGGCT 1963

QY 318 LeuAlaMetLeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAla 337

Db 1964 CTAGCAATGTTGAGACCAACCGCTTCTCTGTGCTGAGCCCTAAATTTATTTAGTCAAGCC 2023

QY 338 GlyValSerSerIleProIleValAspAspAsnAspSerLeuLeuAspThrTyrSerArg 357

Db 2024 CAAAGTAAGCTCAATACCAATAGTTCATGATATGACTCTATTTATGATATATCTGCTCG 2083

QY 358 SerAspIleThrAlaLeuAlaIleCysValTyrThrHisValArgLeuAspGluMet 377

Db 2084 AGTGACATACAGCTTTTGGCAAGAACAGAGCATATACATATTAATCTTTCAGCAATG 2143

QY 378 ThrIleHisGlnAlaLeuGlnIleGlyGlnAspAlaAsnThrProPheGlyPheAsn 397

Db 2144 ACTGTTTCATCAGCTTTGAGTGGCCAGGACGATATAGTCCCTATAG---CTTAGA 2200

QY 398 GlyGlnArgCysGlnMetCysLeuArgSerAspProLeuLeuIleValMetGluArgLeu 417

Db 2201 AGTCAAGATGTCAGATGTTTGGGTTCTGATCTCTGATTAAGTATGAGACGCTTG 2260

QY 418 AlaAsnProGlyValArgValPheIleValGluAlaGlySerLysArgValGluGly 437

Db 2261 GCAATCCAGTGTGAGCGGCTTGTCTATCGTGAAGCTGGCAGCAAGCGTGTAGAGGCG 2320

QY 438 IleIleSerLeuSerAspIlePheIlePheLeuLeu 449

Db 2321 ATGTTTCATGAGTGCATATTCAGTTCTTCATT 2356

RESULT 7

US-10-425-114-6702

; Sequence 6702, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 6702

; LENGTH: 1200

; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700575557_FLI
 ; US-10-425-114-6702

Alignment Scores:

Pred. No.: 5,89e-176 Length: 1200
 Score: 1492.00 Matches: 292
 Percent Similarity: 97.35% Conservatives: 2
 Best Local Similarity: 96.69% Mismatches: 8
 Query Match: 63.71% Indels: 0
 DB: 13 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-425-114-6702 (1-1200)

QY 150 PheHisIleLeuHisGluGlnGlyProValAlaProLeuThrPheSerPheArgGly 169
 DB 2 TTTCATATTCGATGAACAGGGGATTCCTGTAGCTCCTCTCGGGACTCAITCAGAGGT 61
 QY 170 GlnPheValGlyLeuSerProLeuAspPheIleLeuIleLeuArgGluLeuGluThr 189
 DB 62 CAATTGTGCGGCTTCGAGCCATTGGATTTCATCTATATATTCGGGAGCTAGAACT 121
 QY 190 HsGlySerAsnLeuThrGluGlnLeuGluThrHisThrIleSerAlaTrpLysGlu 209
 DB 122 CATGGCTCGAATTCACAGAGATCAGCTTGAACACACACTATATCTGCATGGAAAGAG 181
 QY 210 AlaLysArgGlnThrAsnGlyArgAsnSerGlnTrpArgProGlnGlnHisLeuVal 229
 DB 182 GCTAAGCGCAACTGTGGAGAAATGATGTGTCAGTGGGAGCATCAGCATCTAGTG 241
 QY 230 HisAlaThrProTyGluSerLeuArgAspIleAlaValLysLeuGlnAsnGlyIle 249
 DB 242 CATGCCACCCCTTATGAGTCTCTGAGGACATTCGAGTAAAGCTTTTGCTAATGACAT 301
 QY 250 SerThrValProValIleTrpSerSerSerSerAspGlySerPheProGlnLeuHis 269
 DB 302 TCTACAGTCCCATTTATTTATTCATCATCATCATCATCATCATCATCATCATCATCAT 361
 QY 270 LeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyPheLysAsnSerThrGly 289
 DB 362 CTTCATCCCTTTCTGGAATTTGAATGTATTTTAGTATTTTAAAACTCAACTGTGT 421
 QY 290 AsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrpValProLys 309
 DB 422 AATTGTGCTATTCTGAACCAACCGGTGTGCTCCATTCGGTGGGTTCCTGGTTCGAAA 481
 QY 310 IleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSerSer 329
 DB 482 ATCGGTGATCAACACAGTCTGCTTCATTTGGCTATGTTGCGACCTAATGCACTAGCTCT 541
 QY 330 AlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleValAspAsnAsp 349
 DB 542 SCCCTTAACATGTGTGTTCAAGCTGGAGTGAGTCAATACCAATTTGGATGAAAACGAC 601
 QY 350 SerLeuLeuAspThrTrpSerArgSerAspIleThrAlaLeuAlaLysAspLysValTrp 369
 DB 602 TCCCTGTCTGACACTTACTCTAGAAGTGACATCACAGCCCTAGCTAAAGACAAGTCTAC 661
 QY 370 ThrHisValArgLeuAspGluMetThrIleHsGlnAlaLeuGlnLeuGlyGlnAspAla 389
 DB 662 ACATATGTTGCTTGGATGATGATGACCATTCATCAGGCTTTACAGCTTTGACAAAGATGCC 721
 QY 390 AsnThrProPheGlyPhePheAsnGlyGlnArgCysGlnMetCysLeuArgSerAspPro 409
 DB 722 AATACGCTTTTGGATTTTAAAGGTTCAGAGATGCGAGATGTCCTCCGCTCTGATCCT 781
 QY 410 LeuLeuLysValMetGluArgLeuAlaAsnProGlyValArgValPheIleValGlu 429
 DB 782 TTCTGAAGGTGATGAGCGACTGGGTAAATCCGGGGGTGGGGGTGTTTCATAGTAGAA 841
 QY 430 AlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePheLysPheLeuLeu 449

Db 842 GCTGGAGCAACGTTGGAGGTATTATATCACTAGTGATATTTCAAGTTCITGCTG 901
 QY 450 SerLeu 451
 Db 902 AGCTTG 907
 RESULT 8
 US-10-424-599-79689
 ; Sequence 79689, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 79689
 ; LENGTH: 2315
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_42975C.1
 US-10-424-599-79689

Alignment Scores:

Pred. No.: 7.51e-175 Length: 2315
 Score: 1487.00 Matches: 287
 Percent Similarity: 78.05% Conservatives: 65
 Best Local Similarity: 63.64% Mismatches: 75
 Query Match: 63.49% Indels: 24
 DB: 13 Gaps: 4

US-09-857-525C-2 (1-451) x US-10-424-599-79689 (1-2315)

QY 4 HisLeuProMetSerProIleGluGlyCysProThrValPheGlnAlaIleCysSerLeu 23
 DB 585 CATATAGTATGCTCTCCGATGAGGGATGCCCTGCTGTGTTTCAAGTTGTTGCAACTTA 644
 QY 24 SerProGlyIleHisGluTrpLysPhePheValAspGlyGluTrpArgHisAspGluArg 43
 DB 645 ATGCCGGGGTTTCATCAGTACAAATTAATGTAGATGTGTAGTGGCGGCACCATGACGAG 704
 QY 44 GlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyLeuThrArgGluTy 63
 DB 705 CAGCATTTGTAAATGGGAGCTGTGGAGTAGTGAACTATTTATATATAGTGAAGAACCA 764
 QY 64 AsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspVal 83
 DB 765 GATATCTTACCTTCCATCTTAATAACTGAAACACTGCT---CGATCACATGAGGTT 821
 QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
 DB 822 GACCACTGGAAGCTAATCCAGCATGACTGTGCTGAT----- 860
 QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyLeuAsn 123
 DB 861 -----CTGGAGGTTTCTCGCCACCGTATATCGGTATCTCTGTC 899
 QY 124 LeuHisThrCysTyAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
 DB 900 ACGCATATCTCAATGATGTTTCCGAGTCAAGAAAGGTTCATTCCTTGGATATATAAT 959
 QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
 DB 960 TTACCAAGTTAAGCAAGCATTCCTATGTTCTTACGAAACAGGCTATATCTATGGCTCTCTTA 1019
 QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183

```

1020 TCGGATTTTTCAGGAGCAGTTCGTTGGAGTCTTAGTGCATGAGCTTCATTCATTA 1079
184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThr 203
1080 TTGAAGAGCTGGGGAACATAGTTCAAAATTTGACTCAAGMACAACTCGAGACTCATCT 1139
204 IleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGln----- 221
1140 ATAGCTGCTTGAAGAGGAAATTTCAACAAATTCAGAACACTTCATAGTAATGGGGA 1199
222 -----TrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArg 238
1200 TCATATCCTTGGCGA-----TTTGTTTCATGCGGACCCCATGAATGCTTAATA 1247
239 AspLeuAlaValLysLeuGlnGlnHisGlySerThrValProValIleTyrSerSer 258
1248 GATGCGGCTTTGAAGGTTTTCAGAAACAAAGGTGCAACTGTCCTCATCATATCTTCT 1307
259 SerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLys 278
1308 TCAGAGGATGTTTCATATCCTCACTGCTACATCTTGCTTCCCTATCCGGAATCTAATA 1367
279 CysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProVal 298
1368 GGTATATGAGGCACTTTTAAGCACTCTTGAAGTCTTGAGTCTTGCCCATCTTCAACTCCAGTG 1427
299 CysSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeu 318
1428 GCTCAATACCTTTGGGTACATGATGCTTAGAGTGGGGAACAAATGTCGGCCACTA 1487
319 AlaMetLeuArgProAsnAlaSerLeuSerAlaLeuAsnMetLeuValGlnAlaGly 338
1488 GCAATGTTGATGCCAAGTCTCTCTCGTGTCTGCTTTGCTATGTTTGTTCAGCCGAA 1547
339 ValSerSerIleProIleValAspAsnAspSerLeuLeuLeuAspThrTyrSerArgSer 358
1548 GTTAGCTCAATACCGATGTCGATGCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1607
359 AspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThr 378
1608 GATATTAAGTGGCTTAAAGATAAAGCTTATGCTAGGATATCTCTGATGAAATAGT 1667
379 IleHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGly 398
1668 ATTCAACGAGCAATGCTTTGGGACAGATGCAACTCTCTCTTAGGCTTTACATGCT 1727
399 GlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAla 418
1728 CACAGATGTCACATGCTTTGAGATCTGATTCGCTTTTACAAAGTGTGGAGCGTGGCT 1787
419 AsnProGlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIle 438
1788 AATCCTGGGTAGGAGACTTGCGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCT 1847
439 IleSerLeuSerAspIlePheLysPheLeuLeu 449
1848 ATTTCCTTAAGCATGGTTTCAGATCTTGTGTA 1880

```

RESULT 9

```

US-10-425-114-29527
; Sequence 29527, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B

```

```

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29527
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY057G02_FLI
US-10-425-114-29527

```

```

Alignment Scores:
Pred. No.: 6,99e-155 Length: 1507
Score: 1325.00 Matches: 256
Percent Similarity: 83.24% Conservative: 52
Best Local Similarity: 69.19% Mismatches: 54
Query Match: 56.58% Indels: 8
DB: 13 Gaps: 3

```

US-09-857-525C-2 (1-451) x US-10-425-114-29527 (1-1507)

```

QY 84 AspAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrValSerGluGlyThr 103
DB 3 GATAATGAGCTTTTCGCCGAATGCGCGGTGACGATGGTACTTTGAGTGAGGTGCTG 62
QY 104 LeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSerGluTyrLeuAsn 123
DB 63 CCAAGATATCAGATCTGATGATCAATATCCCGTCAGCGTATTTCTGCAATCTCTCT 122
QY 124 LeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsn 143
DB 123 TCACACACCGCTTATGAATTAATCTCCGAGTCAGCAAGGTGTTGCTTTGATGTTGAT 182
QY 144 LeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeu 163
DB 163 CTACAGTGAAACAGGCAATTCATATATTCATGATGAGGAGGAGTTCATGCTCTCTCT 242
QY 164 TrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuIle 183
DB 243 TGGGACTTCTGCAAGGGGCAATTTGTTGTTCTTCTAGTCTGCTGCTGCTGCTGCTGCT 302
QY 184 LeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThr 203
DB 303 TTAAGAGAGCTGGGGAATCATGATTCATATTCATGATGAGGAGGAGTTCGAAACACATACC 362
QY 204 IleSerAlaTrpLysGluAlaLys-----ArgGlnThrAspGlyArgAsnAsp 219
DB 363 ATATCAGCTTGGAAAGAGGAATCGTATCTAAATAGACAGACAAATGGACATGGAACT 422
QY 220 SerGlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAsp 239
DB 423 GCATTTTCAAGA-----TGTTTATCCATGCGGCGCATATGATAATCTGAAAGAT 473
QY 240 IleAlaValLysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSer 259
DB 474 ATTGCCATGAGATCTTGGAAAGAGGTTTCACTGTTCTTATTCATTCATCTCTCTCT 533
QY 260 SerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCys 279
DB 534 GAAGATGCTTCAATTCACAGTTACTACATCTTGCTTCTCCTTCCAGGAACTTAAATGC 593
QY 280 IleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCys 299
DB 594 ATTGTAGGTATTTTAGGCACTCTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 653
QY 300 SerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAla 319
DB 654 GCAATACCTCTGGGACAGTGGGTGCGCCAAATTTGGGAATCAAAATCGGCGGCTCTAGCA 713
QY 320 MetLeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyVal 339
DB 714 ATGTTGAGACCAACCGCTTCTCTTGGCTCAGCCCTAAATTTATTAGTTCAAGCCCAAGTA 773

```

```

QY 340 SerSerLeuProIleValAspAsnAspSerLeuLeuAspThrTyrSerArgSerAsp 359
DB 774 AGCTCAATACCAATAGTATGATGATAATGACTCATTTATTTGGATATATCTCTGCGAGTGAC 833
QY 360 IleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIle 379
DB 834 ATACAGCTTTGGCAAGACAGACATATACATATTAATCTTTGACGAATGACTGTT 893
QY 380 HisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGln 399
DB 894 CATCAGGCTTTGCGAGTTGGCGCAGGACGATATAGTCCCTATGAG---CTTAGAAGTCAA 950
QY 400 ArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsn 419
DB 951 AGATGTCAGATGTTGCGTCTGATCCTCTGCATAAAGTGATGGAACGCTTGGCAAT 1010
QY 420 ProGlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIle 439
DB 1011 CCAGGTGTCGGCGCTTGTCTATCGTGAAGCTGCGCAGCAAGCGTGTAGAAGGCATTGTT 1070
QY 440 SerLeuSerAspIlePheLysPheLeuLeu 449
DB 1071 TCATTGAGTGACATATTCAAGTTCCTCAT 1100

```

RESULT 10

```

US-10-424-599-105
; Sequence 105, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 105
; LENGTH: 1879
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100095C.1
US-10-424-599-105

```

```

Alignment Scores:
Pred. No.: 1-4e-104 Length: 1879
Score: 924.50 Matches: 196
Percent Similarity: 65.35% Conservative: 36
Best Local Similarity: 55.21% Mismatches: 56
Query Match: 39.47% Indels: 67
DB: 13 Gaps: 6

```

```
US-09-857-525C-2 (1-451) x US-10-424-599-105 (1-1879)
```

```

QY 128 TyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIleAsnLeuProValLys 147
DB 447 TATCGACATTTCCAGATGCGTATAAAGTGCATATTCGATTGTAACCTCATCTATCTT 506
QY 148 GlnSerPheHisIleLeuHisGluGlnGlyIleProValAlaProLeuTrpAspSerPhe 167
DB 507 CAATATTCGGTAATGCTTGTAAACTTGTGTCTT----- 542
QY 168 ArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeuLeuArgGluLeu 187
DB 543 -----TTTCTTTGAAACACAGGCTC 563
QY 188 GluThrHisGlySerAsnLeuThrGluGlnLeuGluThrHisThrIleSerAlaTrp 207
DB 564 CGAAATCATGGTCCCAATCTAACAGAGAGGAGCTTGAACACATACCATATCAGCTTGG 623

```

```

QY 208 LysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgPro---GlnGln 226
DB 624 AAAGCAGGAAAA-----TGGACAGATTTCACAA 653
QY 227 HisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeuLeuGln 246
DB 654 TGTTTATCCGTGCGAGGCCATACGATAATTTGAAAGAGATTGCTGTGAAGATCCTGCAA 713
QY 247 AsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGlySerPheProGln 266
DB 714 CATGAATTTCAACAGTTCTCTATTATTCAT-----TCAGAAGATGGTTCAATTCACAG 767
QY 267 LeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPheLysAsn 286
DB 768 CTACTACATCTGCTTCTACTTTCAGAAATCCTTAATGCAATTTGAGGATTTTATGAAAT 827
QY 287 SerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySerTrp 306
DB 828 TGTTCATGTTTCAATTCCTTATCTTCAACTTCCAAATTTGTGCAATCCCTGTGGCAGCTGG 887
QY 307 ValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAlaSer 326
DB 888 GTGCCCAAAATTTGGGAATCAAAATCGGCAGGCTCTAGCAATGTTGAGACCAACCGCTTCT 947
QY 327 LeuSerSerAlaLeuAsnMetLeuValGln----- 336
DB 948 CTGCGGTGAGCCCTAAATTTATTTAGTTCAGGTGTGAACCTTTTGTCTGACCTTGAACAT 1007
QY 336 ----- 336
DB 1008 AGTGTATCAAAAGAAAGCAACAGAGAGAAGAAATGGAATTTGGTATCCTATTATAC 1067
QY 337 -----AlaGlyValSerSerIleProIleValAspAsnAspSerLeuLeuAspThr 354
DB 1068 TTCTGAGCCCAAGTAAGCTCAATACCAATAGTTCATGATAATGACTATTAATCTTACTGATATA 1127
QY 355 TyrSerArgSerAspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArgLeu 374
DB 1128 TATTTGTCGAGTGCATACACAGCTTTGGCAAGAACAGACAGCATATGCAATTAATCTT 1187
QY 375 AspGluMetThrIleHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProPheGly 394
DB 1188 GATGAATGACAGTTTCATCAGGCTTTGCGAGTGGCGCAGGACGATATAGTCCCTATGAG 1247
QY 395 PhePheAsnGlyGlnArgCysGlnMetCysLeuArgSerAspProLeuLeuLysValMet 414
DB 1248 ---CTTAGAAGTCAAGATGTCAGATGTGTTTGGTTCCTGATCCTCTGCAATAAAGTGATG 1304
QY 415 GluArgLeuAlaAsnProGlyValArgArgValPheIleValGluAlaGlySerLysArg 434
DB 1305 GAACGCTTGGCAAAATCCAGGTGTGAGCGGCTTGTCTCATCGTGAAGCTGCGCAAGCGT 1364
QY 435 ValGluGlyIleIleSerLeuSerAspIlePheLysPheLeuLeu 449
DB 1365 GTAGAAGGCATTGTTTCAATTTGAGGGACATATTCAAGTTCCTTCAAT 1409

```

RESULT 11

```

US-10-424-599-101
; Sequence 101, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 101
; LENGTH: 1225

```


TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1225)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_100091C.1
US-10-424-599-101

Alignment Scores:
Pred. No.: 1.89e-104 Length: 1225
Score: 921.00 Matches: 186
Percent Similarity: 81.56% Conservative: 31
Best Local Similarity: 69.92% Mismatches: 34
Query Match: 39.33% Indels: 15
DB: 13 Gaps: 4

US-09-857-525C-2 (1-451) x US-10-424-599-101 (1-1225)

QY 186 GluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHisThrIleSer 205
DB 21 CAGCTTGAATCATGGTCCCAATCTACAGAGAGAGAGCTTGAACACATACCATCA 80
QY 206 AlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTrpArgPro 224
DB 81 GCTTGGAAAGAGGAGAAA-----TGACAGGATTT 110
QY 225 GlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaValLysLeu 244
DB 111 ACACAACTTTTATCCGTCAGGGCCATACGATAATTTGAAGAGATTGCTGTGAAGATC 170
QY 245 LeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGlySerPhe 264
DB 171 CTGCAACATGGAATTTCAACATGCTCTATTATCAT-----TCGAGAGATGTTCAATT 224
QY 265 ProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArgTyrPhe 284
DB 225 CCACAGCTACTACATCTTGCTTCACTTTCAGGAATCTCTAAATGATCATTTGAGGTATTT 284
QY 285 LysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGly 304
DB 285 AGGAATGTTCTAGTTCAATGCTATATCTTCAACTTCAATTTGCAATCCCTGTGGGC 344
QY 305 SerTrpValProLysIleGlyLeuAsnSerArgProLeuAlaMetLeuArgProAsn 324
DB 345 ACGTGGGTGCCCAAAATTTGGGAGTCAAAATCCCGGCTCTAGCAATGTTGAGACCAAT 404
QY 325 AlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIle 344
DB 405 GCTTCACTTACTTCAGCCCTAAACTATTAGTTCAAGCTCAAGTAAAGTTCAATACCAATA 464
QY 345 ValAspAsnAspSerLeuLeuAsp-ThrTyrSerArgSerAspIleThrAlaLeuAl 364
DB 465 GTTGATGATAGTACTCAATCTAGATATTATTTGTCGAGTGATATAACAGCTTTGGC 524
QY 364 aLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGln 384
DB 525 AAAGACAGAACTTATACCCATATTAATCTTGATGAATGACTGTTCAATCAGGCATTGCA 584
QY 384 nLeuGlyGlnAspAlaAsnThrProPheGlyPheAsnGlyGlnArgCysGlnMetCys 404
DB 585 ATTTGGGCAGAGATTCTTATATACTATTAG---CTGAGTTGTCAAGATGTCAGATGTG 641
QY 404 sLeuArgSerAspProLeuLysValMetGluArgLeuAlaAsnProGlyValArgAr 424
DB 642 TTTGCGAACTGATTCTCTGCAATGAAGTATGAAAGCTTTGGCAATCCAGGTGTGAGCG 701
QY 424 gValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIle 444
DB 702 GCTTGTAAATTTGGAAGCTGCAGCAGCGGTAGAGGCATCATGACATGAGTGACAT 761
QY 444 ePheLysPheLeuLeu 449

DB 762 ATCAACTTCTTCCTT 777

RESULT 12

US-10-437-963-90730
Sequence 90730, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 90730
LENGTH: 6747
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_89371C.1
US-10-437-963-90730

Alignment Scores:
Pred. No.: 1.21e-101 Length: 6747
Score: 908.50 Matches: 203
Percent Similarity: 64.95% Conservative: 36
Best Local Similarity: 55.16% Mismatches: 92
Query Match: 38.79% Indels: 37
DB: 17 Gaps: 9
US-09-857-525C-2 (1-451) x US-10-437-963-90730 (1-6747)

QY 42 GluArgGlnProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArg 61
DB 4912 GAACCGGTACACGCAATTTCCAAACGCTACGGTTGGTAGAGGAGCCCTCTATCGGCC 4971
QY 62 GluTyrAsnGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMet 81
DB 4972 GCGGTAAACGGAATCCTCTGAAGTGCATTCCTCAGGAACAAGGCGTTGAGCTTCTTGT 5031
QY 82 AspValAspAsnGluAsn-----PheGlnArgThrValThrLeuSerAspGlyThr 98
DB 5032 GACATCCATGAGGGGAGTGGGAGCCCATTCGGCTCGCGCACCTGGTTGGCAAGGCC 5091
QY 99 ValSerGluGly-----ThrLeuArgValSerGluAlaIleGlnIleSerArg 115
DB 5092 TTTGCTCAAGGATTTTATTGGCCGACAGCTCTCAATGATGACATCGATCGATCGCGCA 5151
QY 116 CysArgValSerGluTyr-----LeuAsnLeuHis-----ThrCysTyrAspLeu 130
DB 5152 TGCAGAGCGTGTAGTCTCAGCCAGCAATTCATCATCGCGGCCGCCAGCTGCGATC 5211
QY 131 LeuProAspSerGly-----LysValIleAlaLeuAspIleAsnLeuProValLysGln 148
DB 5212 ATACCACTGTATGCGCCATTTGCTGTCTGGGGGCTTGATATCTCTGGGACCGTCAACGG 5272
QY 149 Ser-----PheHisIleLeuHis-----GluGln 156
DB 5272 GCCCGGCGCGGTTTGAGTATCTGTATGTTGCGGTCCGACAAAGTTCAATGAGGCCGAG 5331
QY 157 GlyIleProValAlaProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSer 176
DB 5332 GCTTATCCGGTTGCAAG-----ATCGATAAGCACTCT 5364
QY 177 ProLeuAspPheIleLeuLeuArgGlu-----LeuGluThrHisGlySerAsn 193

```

Db      5365  GCTCTAAATTCATTAAGGCGCATCAGCGCGTTTGGACTGGAAATCATGGCTCCAAT 5424
Qy      194  LeuThrGluGlnLeuGluThrHisThrIleSerAlaTrpLysGluAlaLysArgGln 213
Db      5425  CTGACAGAAGACGACGCTTGAACACATACCTATATCTGGTGAAGAGGCGCAAGCGTCAA 5484
Qy      214  ThrAsnGlyArgAsnAspSerGlnTrpArgProGlnHisLeuValHisAlaThrPro 233
Db      5485  ACTATGCCGAACAGAGGGTCTTGGAGGCAATCACCATTAGTTCTGCGCACCCCT 5544
Qy      234  TyrGluSerLeuArgAspIleAlaValLysLeuGlnHisLeuHisLeuAlaSerLeu 253
Db      5545  TATGAATCCCTGAGCGAAATGCTATGAGACTCTGCAAAACGGTGTCTTACCGTTCCA 5604
Qy      254  ValIleTyrSerSerSerAspGlySerPheProGlnLeuLeuHisLeuAlaSerLeu 273
Db      5605  ATTATGTTTTCCTCATCACAGATGGCTCATATCCCAATTTGTCATCTTGTTCCTT 5664
Qy      274  SerGlyIleLeuLysCysIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIle 293
Db      5665  TCTGGAATTTGAATGATTTTGTAGATACTTAAATAATCTCAAGGTAAATTTACCTATT 5724
Qy      294  LeuAsnGlnProValCysSerIleProLeuGlySerTrpValProLysIleGlyAspLeu 313
Db      5725  TTGAGCCAACTGATGACAAATTCCTCTGGGTACCTGGGTTCACAAAATTTGGTGATCCT 5784
Qy      314  AsnSerArgProLeuAlaMetLeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMet 333
Db      5785  AATGGCGCTCATGGCTATGTCGGGCTTAACACATCTCTTAGCGTGCCTCGAATTTG 5844
Qy      334  LeuValGlnAlaGlyValSerIleProIleValAspAspAsnAspSerLeuLeuAsp 353
Db      5845  CTGTTCAAGCTGGTGTGAGCTCAATCAATCAATGTTGATGATGATGATGATGATGAT 5904
Qy      354  ThrTyrSerArgSerAspIleThrAlaLeuAlaLysAspLysValTyrThrHisValArg 373
Db      5905  ACATATTCCAGAAGTGACATCACAGCTCTTGCAAAAGACAAAGGTTGACACACATTCGC 5964
Qy      374  LeuAspGluMetThrIleHisGln 381
Db      5965  CTAGATGAGATGACCATTCATCAG 5988

```

```

RESULT 13
US-10-437-963-66100
; Sequence 66100, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bouharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 66100
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_67084C.1
US-10-437-963-66100

```

Alignment Scores:
Pred. No.: 6,75e-92 Length: 1402
Score: 821.50 Matches: 179

```

Percent Similarity: 58.85% Conservative: 87
Best Local Similarity: 39.60% Mismatches: 36
Query Match: 35.08% Indels: 36
DB: 17 Gaps: 6

US-09-857-525C-2 (1-451) x US-10-437-963-66100 (1-1402)

Qy      10  IleGluGly-CysPro-----ThrValPheGlnAlaIleCysSerLeuSe 24
Db      69  GTGAGGAGGAGTGGCCCATATGGGGCTCGTGGGGCCGAGTTCCAGGTGGTTCGATCTGCC 128
Qy      24  rProGlyIleHisGluTyrLysPhePheValAspGlyGluTrpArgHisAspGluArgG1 44
Db      129  CCGCGGGGTTTATCAGTACCGGTTTGGTGTGATGCTGTGGCGATGTGATGACAGCA 188
Qy      44  nProThrIleSerGlyGluPheGlyIleValAsnThrLeuTyrLeuThrArgGluTyrAs 64
Db      189  ACCTTGTCACGGGATGAATATGGGTTGATTAGCAATGAAGTCTCTGTGGATAATACACA 248
Qy      64  nGlnIleAsnThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspValAs 84
Db      249  TCCGGTT-----GTACA 260
Qy      84  pAsnGluAsnPheGlnArgThrValThrLeuSerAspGlyThrVal----- 99
Db      261  GCCAGAACTTCCATCAGAGTAGTAGTATGGATGAGGTACTACTATGACACAAATGCC 320
Qy      100  -SerGluGlyThrLeuArgValSerGluAlaAlaIleGlnIleSerArgCysArgValSe 119
Db      321  CCCAGATCAGTTATCTCAAAACTCAGGCGTGCAGATAGCAATTTTCCGCCATCGAGTCTC 380
Qy      119  rGluTyrLeuAsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAl 139
Db      381  TGAGATATTATCATATACCATATATGATGTTGTTCTCTGTTCTACAGATAGTACAGT 440
Qy      139  aLeuAspIleAsnLeuProValLysGlnSerPheHisIleLeuHisGluGlnGlyIlePr 159
Db      441  TTTGATGCTCGTCTCTCTCAAGCAAGCCCTTTAAATCATGATGATGAGGGTCTTTC 500
Qy      159  oValAlaProLeuTrpAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAs 179
Db      501  TTTGGTTCCTCTTTGGGATGATCAGCAGCAACCCGTAACAGGCATGCTAACCCGATCCGA 560
Qy      179  pPheIleLeuLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGlnLe 199
Db      561  TTTTGTATTAATTTTGAGAAAGCTGCAGAGAAACATTCGAACTCTTGGCCATGAGAACT 620
Qy      199  uGluThrHisThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAs 219
Db      621  TGAATGCAATTCAGTGTCTCTTGAAGAAGCAAAATTCAGTATTATGGGGACCTGA 680
Qy      219  pSerGlnTrpArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAs 239
Db      681  TGTGTGCCATTGAGAGAGCCATTAATTCATGTTAAGGATTCAGATAATTTAAGGGA 740
Qy      239  pIleAlaValLysLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSe 259
Db      741  TGTGGCATTCGCTATATAGAAATGAATATCTCAGTTCTTCCCTTTCAGTTCCTT 797
Qy      259  rSerAspGlySerPheProGlnLeuHisLeuAlaSerLeuSerGlyIleLeuLysCy 279
Db      798  AAGGATTCATCAGGGATGCTTTGCTTGTCTTGCACCCCTTCCAGGGATTTGAAAT 857
Qy      279  sIleCysArgTyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCy 299
Db      858  TATTGCTCAAGCTCAAGACAGCCTGAAGGTACTATTCTTCGAAAACCCAGATTGT 917
Qy      299  sSerIleProLeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAl 319
Db      918  CAGTATGCCCTATTGGTACATGTCACCATCTATCGCAAGCGCAAGCAATAGACAGCTAG 977
Qy      319  aMetLeuArgProAsnAlaSerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyVa 339

```

Db 978 AACTTCGGCGACCAAGCACTCTCTAAATTCAAGCTGATTTGCTTGAGATAGAGT 1037
 Qy 339 lserSerIleProIleValAspAsnAspSerLeuLeuAspThrTySerArgSerAs 359
 Db 1038 AAGCTCAATTCCTATAGTTGACGATAATGGCGCTCTCTTGATGCTACTCGCTCAGTGA 1097
 Qy 359 pIleThrAlaLeuAlaLysAspLysValTyThrHisValArgLeuAspGluMetThrIl 379
 Db 1098 TATCATGGCTCTTAGCAAGAATGAGTCTACACTCGTATTGAGCTTGAAACAGGTACCGGT 1157
 Qy 379 eHisGlnAlaLeuGlnLeuGlyGlnAspAlaAsnThrProp:eGlyPhePheAsnGlyG1 399
 Db 1158 GGAGCATGCCITGGAGCTGCAATACCAAGTG-----AATGGCGG 1196
 Qy 399 nArg---CysGlnMetCysLeuLeuArgSerAspProLeuLeuLysValMetGluArgLeuAl 418
 Db 1197 AAGACACTGTCTATACCTCTTGAGCACTAGTACCTCTCTGGAGGTTTGGAGCAATTGTC 1256
 Qy 418 aAsnProGlyValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyTl 438
 Db 1257 AGCTCCAGGGGTGGCGGAGTGGTCTGTTATTGAACCAAGGAGCAGATTGTGCAAGGAAT 1316
 Qy 438 eIleSerLeuSerAspIlePheLysPheLeuLeu 449
 Db 1317 AATCTCATGAGGAGCGCATTTACATTCTCATTT 1350

RESULT 14

US-10-437-963-22833
 ; Sequence 22833, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kowalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Audrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 22833
 ; LENGTH: 1512
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_27972C.1
 US-10-437-963-22833

Alignment Scores:
 Pred. No.: 4,06e-85 Length: 1112
 Score: 766.00 Matches: 152
 Percent Similarity: 95.78% Conservative: 7
 Best Local Similarity: 91.57% Mismatches: 6
 Query Match: 32.71% Indels: 1
 DB: 17 Gaps: 0

US-09-857-525C-2 (1-451) x US-10-437-963-22833 (1-1112)

Qy 286 AsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIleProLeuGlySer 305
 Db 5 AATCTCAAGGTAATTTACCTATTTTGAGCCAAAGCTGTATGCACAAATTCCTCGGTACC 64
 Qy 306 TrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArgProAsnAla 325
 Db 65 TGGGTTTCAAAATAATGGTGATTCCTAATGGCCGTTCATTGGCTATGTTGGCCCTAACACA 124
 Qy 326 SerLeuSerSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIleProIleVal 345

Db 125 TCTCTTAGGGCTGCCTCGAATTGCTGGTCAAGCTGGTGTGAGCTCAATACCAATTGTG 184
 Qy 346 AspAspAsnAspSerLeuLeuAspThrTySerArgSerAspIleThrAlaLeuAlaLys 365
 Db 185 GATGACAATGACTCACTGCTTGACACATATTCACAGAAGTGACATCACAGCTCTTGCAAAA 244
 Qy 366 AspLysValTyThrHisValArgLeuAspGluMetThrIleHisGlnAlaLeuGlnLeu 385
 Db 245 GACAAGGTGTACACACATTCGCTAGATGAGATGACCAATTCATCAGGCTCTGCAGCTT 304
 Qy 386 GlyGlnAspAlaAsnThr-PropHeGlyPhePheAsnGlyGlnArgCysGlnMetCysLe 405
 Db 305 GCACAGGATGCAAATTCGCCCTTTGGATTCTTTAAACGGGCAAGATGTCAGATGTGCT 364
 Qy 405 nArgSerAspProLeuLeuLysValMe-Si:nArgLeuAlaAsnProGlyValArgArgVa 425
 Db 365 CCGGTCTGACACTTTGTTGAAAGTGATGAGGCGATTAGCTAATCTCGGGGTGGCGGTGT 424
 Qy 425 lPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeuSerAspIlePh 445
 Db 425 TTTCAATTGTGAAGCTGTGTAGCAAAACGTGTGGAGGGCATAATATCACTCAGTGATATTTT 484
 Qy 445 elysPheLeuLeuSer 450
 Db 485 CAAGTTCTTGCTGAGC 500

RESULT 15

US-10-425-114-8857
 ; Sequence 8857, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kowalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 8857
 ; LENGTH: 1533
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700799332_FLI
 US-10-425-114-8857

Alignment Scores:
 Pred. No.: 2,94e-84 Length: 1533
 Score: 761.00 Matches: 173
 Percent Similarity: 57.59% Conservative: 85
 Best Local Similarity: 38.62% Mismatches: 152
 Query Match: 32.49% Indels: 28
 DB: 13 Gaps: 8

US-09-857-525C-2 (1-451) x US-10-425-114-8857 (1-1533)

Qy 13 CysPro-----ThrValPheGlnAlaIleCysSerLeuSerProGlyIle 27
 Db 168 TGTCCCATGGGCTGGTGGGCGCGTGTTCAGGTGCTTCGATCTGCCCCCGGAGTT 227
 Qy 28 HisGluTyrlsPhePheValAspGlyGluTrpArgHisAspGluArgGlnProThrIle 47
 Db 228 TACCAGTACCGGTTTTGGTTGATGGTCTCTGGAGGTGTGATGAGACAAAACCCCTTTGTA 287
 Qy 48 SerGlyGluPheGlyIleValAsnThrLeuTyrlleuThrArgGluTyrlAsnGlnIleAsn 67
 Db 288 CGTGATGAATATGAGTCAATGATCAGCAATGAGTGTCTGTGGAAAACAATGTACACCTGTT 347

QY 68 ThrLeuSerSerProSerThrProGlySerArgMetAsnMetAspValAspSerGluAsn 87
Db 348 GTGCGACGAGCCCTTCTATCAGAGAACT- - - - - AATATGGAT- - - - - SerGluGly 102
QY 88 PheGlnArgThrValThrLeuSerAspGlyThrVal- - - - - SerGluGly 102
Db 387 - - - - - AGGGTACCATTTTGAAACAATGCCCGGAGCCA 422
QY 103 ThrLeuArgValSerGluAlaIleGlnIleSerArgCysArgValSerGluTyrLeu 122
Db 423 TCATCTCAGAACCCAGCATGCAATAGCAGTTATCGCCATGTGGTCTCTGGAATATTA 482
QY 123 AsnLeuHisThrCysTyrAspLeuLeuProAspSerGlyLysValIleAlaLeuAspIle 142
Db 483 TTACACAATACCATATATGAGTGTTCCTCCCTTCTAGCAAGTTAACAGTTTGGACACT 542
QY 143 AsnLeuProValLysGlnSerPheHisIleLeuHisIleGlnGlyIleProValAlaPro 162
Db 543 CAGCTTCTCTGTTAAACAAGCATTTAAATAATGATGAGGGTCTTGTCTGGTTCCT 602
QY 163 LeuTyrAspSerPheArgGlyGlnPheValGlyLeuLeuSerProLeuAspPheIleLeu 182
Db 603 CTTTGGGATCAGCGTCAGGAGCAACATAACAGGCATGCTCATCATCAGATTGTGTTTA 662
QY 183 IleLeuArgGluLeuGluThrHisGlySerAsnLeuThrGluGluGlnLeuGluThrHis 202
Db 663 ATCTTGAGAAAGTTGCAAGAAACATTCAGATTATTGGCAACGAAGAG- - - - - 710
QY 203 ThrIleSerAlaTrpLysGluAlaLysArgGlnThrAsnGlyArgAsnAspSerGlnTyr 222
Db 711 CCCATTTCCTGTTGGAAGAAGCAAGCTACAGTTC-ATGGTGGGCGCTGATGGTGGCC 770
QY 223 ArgProGlnGlnHisLeuValHisAlaThrProTyrGluSerLeuArgAspIleAlaVal 242
Db 771 ATCCAGAGAGGCCATTATCCATGTTAAGGATTCAGATTAATTTAGTGATGGCACTG 830
QY 243 LysLeuLeuGlnAsnGlyIleSerThrValProValIleTyrSerSerSerSerAspGly 262
Db 831 ACTATAATCAGAAATGAATATCTTCAGTTCTCT-ATCTTTAAGTCATGGCAGATTCA 887
QY 263 SerPheProGlnLeuLeuHisLeuAlaSerLeuSerGlyIleLeuLysCysIleCysArg 282
Db 888 TCAGGGGTGCTTCTCTTAATCTTGCAACCTTACAGGGGATTTGAAATTTCTTTGCTCG 947
QY 283 TyrPheLysAsnSerThrGlyAsnLeuProIleLeuAsnGlnProValCysSerIlePro 302
Db 948 AAGCTACAGAGAGGCTGAGGGCTGTTCCTTCCTGCACAATCAGCTTCTCAGATTCTCT 1007
QY 303 LeuGlySerTrpValProLysIleGlyAspLeuAsnSerArgProLeuAlaMetLeuArg 322
Db 1008 ATTGGTACATGGTCTCCACATACGGGGAGGTCAAGTAGCAGGCAACTCAGAACTTTGCTA 1067
QY 323 ProAsnAlaSerLeuSerAlaLeuAsnMetLeuValGlnAlaGlyValSerSerIle 342
Db 1068 CTGAGTTCCTCTAATACCTGCTGGAATTCCTGCTTCAAGATAGATAGCTGAT 1127
QY 343 ProIleValAspAspAsnSerLeuLeuAspThrTyrSerArgSerAspIleThrAla 362
Db 1128 CCTATAGTTGATGACAATGATCCCTCGGTGATGCTTACTCTCAGTGATATCATGGCT 1187
QY 363 LeuAlaLysAspLysValTyrThrHisValArgLeuAspGluMetThrIleHisGlnAla 382
Db 1188 CTGGCAAGAAGTATGTTATGCTCGCATCGAACTTGAAACAGTGCACCGTACAAAATGCT 1247
QY 383 LeuGlnLeuGlyGlnAspAlaAsnThrProPheGlyPhePheAsnGlyGlnArg- - - Cys 401
Db 1248 TTGGATGTGCATACAGGTG- - - - - CATGCCGCAAGACAGTGT 1286
QY 402 GlnMetCysLeuArgSerAspProLeuLeuLysValMetGluArgLeuAlaAsnProGly 421
Db 1287 CATACTTGTTTACAGCAGTACCTTGTCTGGAAGTTTGGAGGGATTGTCATTCACAGGA 1346
QY 422 ValArgArgValPheIleValGluAlaGlySerLysArgValGluGlyIleIleSerLeu 441

Db 1347 GTGCGACGCTTGTGTTATTGAACAAAGTACCAGATTTTGTGGAAGGAATCATCTCATTTG 1406
QY 442 SerAspIlePheLysPheLeuLeu 449
Db 1407 AGAGACGTATTACATTCTCCCTT 1430
Search completed: July 9, 2004, 11:24:30
Job time : 697 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: July 7, 2004, 17:48:03 ; Search time 45 Seconds
(without alignments)

3162.194 Million cell updates/sec

Title: US-09-857-525C-2
Perfect score: 2342
Sequence: 1 TREHLPMSIEGPTVFQAL.....SKRVEGIIISLDIFKLLSL 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2315	98.8	496	10 Q9FUY4	Q9fuy4 zea mays (m
2	2218	94.7	497	10 Q9FUY5	Q9fuy5 zea mays (m
3	2085	89.0	493	10 Q84M80	Q84m80 oryza sativ
4	1580	67.5	487	10 Q944A6	Q944a6 arabidopsis
5	1481.5	63.3	485	10 Q84PE0	Q84pe0 medicago tr
6	1334.5	57.0	382	10 Q9FV59	Q9fv59 arabidopsis
7	1030.5	44.0	391	10 C04028	C04028 arabidopsis
8	820.5	35.0	451	10 Q7XV93	Q7xv93 oryza sativ
9	405	17.3	577	5 Q86KN7	Q86kn7 dictyosteli
10	376.5	16.1	443	11 Q8BNQ9	Q8bnq9 mus musculu
11	374	16.0	566	11 Q91WG5	Q91wg5 mus musculu
12	369	15.8	585	5 Q8IN38	Q8in38 drosophila
13	365	15.6	906	5 Q8SXT8	Q8sxt8 drosophila
14	365	15.6	582	5 Q9BHL6	Q9bhl6 caenorhabdi
15	363.5	15.5	647	5 Q966I3	Q966i3 drosophila
16	363.5	15.5	906	5 Q810B2	Q810b2 drosophila

17	363	15.5	718	5	Q810K6	Q810k6 drosophila
18	363	15.5	814	5	Q810A1	Q810a1 drosophila
19	361.5	15.4	947	5	Q9VDD6	Q9vdd6 drosophila
20	361.5	15.4	1400	5	Q9VDD2	Q9vdd2 drosophila
21	352.5	15.1	340	4	Q8N7V9	Q8n7v9 homo sapien
22	350	14.9	538	5	Q81GX0	Q81gx0 drosophila
23	333.5	14.2	337	13	Q7ZYL2	Q7zyl2 xenopus lae
24	330.5	14.1	464	11	Q8CJ41	Q8cj41 mus musculu
25	330.5	14.1	489	11	Q8BGM7	Q8bgm7 mus musculu
26	324	13.8	490	11	Q80WK8	Q80wk8 mus musculu
27	287.5	12.3	267	11	Q8BM63	Q8bm63 mus musculu
28	228	9.7	423	5	Q02168	Q02168 caenorhabdi
29	217	9.3	421	5	Q8IN37	Q8in37 drosophila
30	201	8.6	439	5	Q9NS01	Q9ns01 caenorhabdi
31	163.5	7.0	488	5	Q81112	Q81112 caenorhabdi
32	154	6.6	460	5	Q22022	Q22022 caenorhabdi
33	151.5	6.5	855	16	Q7UK82	Q7uk82 rhodopirell
34	137.5	5.9	465	5	Q8MPT5	Q8mpt5 caenorhabdi
35	129	5.5	63	11	Q925V0	Q925v0 mus musculu
36	123	5.3	304	5	Q8T849	Q8t849 dictyosteli
37	121	5.2	271	13	Q7S297	Q7s297 xenopus lae
38	121	5.2	393	17	Q8TZ50	Q8tz50 methanopyru
39	117.5	5.0	333	17	Q8TH73	Q8th73 methanosarc
40	117	5.0	335	5	Q86J96	Q86j96 dictyosteli
41	115	4.9	629	3	Q94035	Q94035 candida alb
42	115	4.9	726	1	Q9P9A0	Q9p9a0 thermococu
43	114	4.9	172	16	Q8EIN2	Q8ein2 shewanella
44	114	4.9	399	10	Q9SSD0	Q9ssd0 arabidopsis
45	113	4.8	357	10	Q8GX19	Q8gx19 arabidopsis

ALIGNMENTS

RESULT 1

Q9FUY4	PRELIMINARY;	PRT;	496 AA.
ID	Q9FUY4		
AC	Q9FUY4;		
DT	01-MAR-2001 (TrEMBLrel. 16, Created)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Protein kinase AKINbetagamma-2.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	PACCAD clade; Panicoideae; Andropogoneae; Zea.		
OX	NCBI_TaxID=4577;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Victoria L., Mar A., Tatjana K., Csaba K., Montserrat P.;		
RT	"Domain fusion between Snf1 related kinase subunits during plant evolution."		
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF276086; ARG31752.1; -		
DR	GO; GO:0016301; F:Kinase activity; IEA.		
DR	InterPro; IPR00644; CBS_domain.		
DR	Pfam; PF00571; CBS; 3.		
DR	SMART; SM00116; CBS; 3.		
KX	Kinase.		
SQ	SEQUENCE 496 AA; 54799 MW; 316P8282B3B55D8A CRC64;		

Query Match 98.8%; Score 2315; DB 10; Length 496;
Best Local Similarity 93.3%; Pred. No. 9.9e-188;
Matches 446; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY	3	EHLPMSPIEGPTVFQALCSLSPGIEHYKFFVDGEWHDROPTISGFGVNTLYLTRE	62
DB	48	EHLPMSPIEGPTVFQALCSLSPGIEHYKFFVDGEWHDROPTISGFGVNTLYLTRE	107
QY	63	YNQINTLSSPTGSRMMDVDNENFQRTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL	122
DB	108	YNQINTLSSPTGSRMMDVDNENFQRTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL	167


```

QY 123 NLHCTCYDLLPDSGKVTALDINKLVKQSFHILHEOGIPVAPLWDSRFGQFVGLLSPDLDFIL 182
DB :|||||
QY 166 SLHCTCYDLLPDSGKVIADINKLVKQSFHILHEOGIPVAPLWDSRFGQFVGLLSPDLDFIL 227
DB :|||||
QY 183 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 242
DB :|||||
QY 243 KLLONGISTVPVIYSSSDGSPQLHLASLSGLKICICRYFNKSTGNLPIILNQPVCSIP 302
DB :|||||
QY 288 KLLONGISTVPVIYSSSDGSPQLHLASLSGLKICICRYFNKSTGNLPIILNQPVCSIP 347
QY 303 LGSWPKIGDLNSRPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSLDDTYSSRSDITA 362
DB :|||||
QY 363 LAKDKVYTHVRLEMTIHOALQLGQDANTFFGFNGQRCQMCCLRSDDLKVMERLANPGV 422
DB :|||||
QY 408 LAKAKVYTHVRLEMTIHOALQLGQDANTFFGFNGQRCQMCCLRSDDLKVMERLANPGV 467
QY 423 RRVFIVEAGSKRVEGIIISLSDIFKFLSL 451
DB :|||||
QY 468 RRVFIVEAGSKRVEGIIISLSDIFKFLSL 496

RESULT 2
Q9FUY5 PRELIMINARY; PRT; 497 AA.
AC Q9FUY5;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DE Hypothetical protein AKIN26gamma-1.
CS Zea mays (Maize);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
CX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Victoria L., Mar A., Tatjana K., Csaba K., Montserrat P.;
RT "Domain fusion between Snf1 related kinase subunits during plant
RT evolution.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF276085; AAG31751.1; -
DR GO; GO:0016301; P:kinase activity; IEA.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 3.
DR SMART; SM00116; CBS; 3.
KW Kinase.
SQ SEQUENCE 497 AA; 54938 MW; B9EF6A530CIAD8B CRC64;

Query Match 94.7%; Score 2218; DB 10; Length 497;
Best Local Similarity 95.1%; Pred. No. 1.7e-179;
Matches 427; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 3 EHLPMSPTEGCTVFOAICSLSPGHEHYKFFVDSGEWRHDEROPTISGFGVNTLYLTRE 62
DB :|||||
QY 49 EHLPMSPTEGCTVFOAICSLSPGHEHYKFFVDSGEWRHDEROPTISGFGVNTLYLTRE 108
DB :|||||
QY 63 YNQINTLSSPTGSRMMMDVDNENFQTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL 122
DB :|||||
QY 109 FNQINALLNPSTGSRMMMDVDNENFQTVTLSDGTIPEGVVRVSEAAIQISRCRVSEYL 168
QY 123 NLHCTCYDLLPDSGKVIADINKLVKQSFHILHEOGIPVAPLWDSRFGQFVGLLSPDLDFIL 182
DB :|||||
QY 169 NLHCTCYDLLPDSGKVIADINKLVKQSFHILHEOGIPVAPLWDSRFGQFVGLLSPDLDFIL 228
QY 183 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 242
DB :|||||
QY 229 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 288
QY 243 KLLONGISTVPVIYSSSDGSPQLHLASLSGLKICICRYFNKSTGNLPIILNQPVCSIP 302

```

```

DB :|||||
QY 289 KLLINDISTVPVIYSSSDGSPQLHLASLSGLKICIFRYFNKSTGNLPIILNQPVCSIP 348
DB :|||||
QY 303 LGSWPKIGDLNSRPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSLDDTYSSRSDITA 362
DB :|||||
QY 349 LGSWPKIGDLNSRPLAMLRPNASLSSALNMLVQAGVSSIPVDDNDSLDDTYSSRSDITA 408
QY 363 LAKDKVYTHVRLEMTIHOALQLGQDANTFFGFNGQRCQMCCLRSDDLKVMERLANPGV 422
DB :|||||
QY 409 LAKAKVYTHVRLEMTIHOALQLGQDANTFFGFNGQRCQMCCLRSDDLKVMERLANPGV 468
QY 423 RRVFIVEAGSKRVEGIIISLSDIFKFLSL 451
DB :|||||
QY 469 RRVFIVEAGSKRVEGIIISLSDIFKFLSL 497

RESULT 3
Q84M80 PRELIMINARY; PRT; 493 AA.
AC Q84M80;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DE Hypothetical protein OSJNBa0059G06.25.
CS OSJNBa0059G06.25.
OC Oryza sativa (japonica cultivar-group);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
CX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zsammann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBa0059G06 genomic sequence.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Buell R.i
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC096690; AAP21389.1; -
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 3.
DR SMART; SM00116; CBS; 3.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 54493 MW; E0531CEA7481CD5F CRC64;

Query Match 89.0%; Score 2085; DB 10; Length 493;
Best Local Similarity 87.3%; Pred. No. 3.3e-168;
Matches 391; Conservative 34; Mismatches 23; Indels 0; Gaps 0;

QY 3 EHLPMSPTEGCTVFOAICSLSPGHEHYKFFVDSGEWRHDEROPTISGFGVNTLYLTRE 62
DB :|||||
QY 46 EHLPMSPTEGCTVFOAICSLSPGHEHYKFFVDSGEWRHDEROPTITGTYGVNTLCITRD 105
QY 63 YNQINTLSSPTGSRMMMDVDNENFQTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL 122
DB :|||||
QY 106 FQINTILSPSTGSRMMMDVDNENFQTVTLSDGTIPEGVVRVSEAAIQISRCRVSEYL 165
QY 123 NLHCTCYDLLPDSGKVIADINKLVKQSFHILHEOGIPVAPLWDSRFGQFVGLLSPDLDFIL 182
DB :|||||
QY 166 NGQGYDLLPDSGKVIADINKLVKQSFHILHEOGIPVAPLWDSRFGQFVGLLSPDLDFIL 225
QY 183 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 242
DB :|||||
QY 226 ILRELETHGNSLITEQLETHTSANKEAKRQTNCRNDSQWRPQOHLVHATPYESLRDIIV 285

```

```

QY 243 KLLONGISTVPVTVSSSDGSPOLLHLASLSGLKICRYFKNSTGNLPLINOPVCSIP 302
DB 286 KILONGVSTVPVIFSSSPDQSYQOLLHLASLSGLKICRYFKNSQGNLPLISQPVCTIP 345
QY 303 LGSWVPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNSLDDTYSRSDITA 362
DB 346 LGTWVPKIGDPNGRPLAMLRPNASLSALNMLVQAGVSSIPVDDNSLDDTYSRSDITA 405
QY 363 LAKDKVYTHVRLDEMTTHQALQGDANTPGFPNGORCOMCLRSDDLKVMERLANPGV 422
DB 406 LAKDKVYTHVRLDEMTTHQALQGDANSPPFPNGORCOMCLRSDDLKVMERLANPGV 465
QY 423 RRVFIVEAGSKRVEGIISLSDIFKLLS 450
DB 466 RRVFIVEAGSKRVEGIISLSDIFKLLS 493

RESULT 4
Q944A6 PRELIMINARY; PRT; 487 AA.
AC Q944A6
DC 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE At1g09020/f7g19.11.
OS Arabidopsis thaliana (Mouse-ear cress).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
CC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
CX NCBI_TaxID=3702;
RN [1]
RN [2]
SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesza E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.X., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA "Arabidopsis cDNA clones";
RL Submitted (0CT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA Chen R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan J.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA "Arabidopsis ORF clones";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF439826; AAL27498.1; -.
DR EMBL; BT000625; AAN18191.1; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 3.
SQ SEQUENCE 487 AA; 53466 MW; 4545FE3BF2C4EBEC CRC64;

```

```

Query Match 67.5%; Score 1580; DB 10; Length 487;
Best Local Similarity 67.6%; Pred. No. 2.4e-125;
Matches 304; Conservative 64; Mismatches 74; Indels 8; Gaps 5;

QY 3 EHLPMSPIECPTVFQAIQCSLSGIIHYKFFVGDGWRHDEROQTISGFGIVNTLYLTRE 62
DB 45 EHVPMSPIECPTVFQVQICNLTPGYHQYKFFVGDGWRHDEROQTFVSGGGVNTITGP 104
QY 63 YNOINTLSSPTSGSRMNMVDNENFORTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL 122
DB 105 -DNVPAGFPETLG-RSNMVDV--DVFLRTADPS-----QEAVPRMSGVDLCLSRHSVLL 157

```

```

QY 123 NLHTCYDLPDQSKVIALDINLPVKQSPHILHEGIPWAPLWDSPRGQVGLLSPLDFIL 182
DB 158 STRTAYELLPSGKVIALDINLPVKQAFPHILYEGIPLAPLWDSPRGQVGLVGLDFIL 217
QY 183 ILRELETHGSLNLTBEQLETHITSAWKEAKQTNCRNDSQWRP-QQHLVHATPYESLRDIA 241
DB 218 ILRELTHGSLNLTBEQLETHITAAWKEGKAHISQYDGSGRFPYPRPLVQVQPYDNLKQVA 277
QY 242 VKLLONGISTVPVTVSSSDGSPOLLHLASLSGLKICRYFKNSTGNLPLINOPVCSIP 301
DB 278 LKILQNVAAVFPVTVSSYQDGSYPOLLHLASLSGLKICRYFKHSSSLPLIQOPICSI 337
QY 302 PLGSWVPKIGDLSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNSLDDTYSRSDIT 361
DB 338 PLGTWVPKIGESSKPLATLRPHASLSALALLVQAEVSSIPVDDNSLDDTYSRSDIT 397
QY 362 ALAKDKVYTHVRLDEMTTHQALQGDANTPGFPNGORCOMCLRSDDLKVMERLANPG 421
DB 398 ALAKDKYATQIHLDDMTVHQAOLQGDASPPYGFNGORCHMCLRSDSLAVKMERLANPG 457
QY 422 VRRVFIIVEAGSKRVEGIISLSDIFKLLS 451
DB 458 VRRVFIIVEAGSKRVEGIISLSDIFKLLS 487

RESULT 5
Q84PEO PRELIMINARY; PRT; 485 AA.
AC Q84PEO
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AKIN betagamma.
OS Medicago truncatula (Barrel medic).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
CC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
CX NCBI_TaxID=3880;
RN [1]
RN [2]
SEQUENCE FROM N.A.
RA STRAIN=cv. Paraggio;
RA Buitink J., Thomas M., Gissot L., Leprince O.;
RA "Differential expression patterns of beta and gamma subunit isoforms
RA suggest the involvement of different SnRK1 complexes in desiccation
RA tolerance, osmotic stress and starvation in germinating seeds of
RA Medicago truncatula.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY247268; AAO61673.1; -.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 3.
DR SMART; SM00116; CBS; 3.
SQ SEQUENCE 485 AA; 53607 MW; 18E908AEBE916F48 CRC64;

```

```

Query Match 63.3%; Score 1481.5; DB 10; Length 485;
Best Local Similarity 64.1%; Pred. No. 5.4e-117;
Matches 287; Conservative 65; Mismatches 83; Indels 13; Gaps 4;

QY 3 EHLPMSPIECPTVFQAIQCSLSGIIHYKFFVGDGWRHDEROQTISGFGIVNTLYLTRE 62
DB 49 EHVPMSPMEGCPVQVQICNLTPGYHQYKFFVGDGWRHDEROQTFVSGGGVNTIYLVRE 108
QY 63 YNOINTLSSPTSGSRMNMVDNENFORTVTLSDGTVSEGLRVSEAAIQISRCRVSEYL 122
DB 109 PDILPAISAEI--SSRSHMEVDNDVFGH-----AANRMSPSDLEVSRRISKFL 158
QY 123 NLHTCYDLPDQSKVIALDINLPVKQSPHILHEGIPWAPLWDSPRGQVGLLSPLDFIL 182
DB 159 SEHTAYDLPDQSKVIALDINLPVKQAFPHILYEGIPLAPLWDSPRGQVGLVGLDFIL 218
QY 183 ILRELETHGSLNLTBEQLETHITSAWKEAKQTNCRNDSQWRP-QQHLVHATPYESLRDIA 241
DB 219 ILKELGTHGSHLTBEQLETHITAAWKEGKSKORRALDNESNPHCFVHAGPKCLKQVA 278

```

```

QY 242 VKLLQNGISTVPVIYSSSDGSPQLLHLASLGILKICICRYFNSTGNLPIINQPVCSI 301
D 279 LKVLQKQKSTVPRI--SLDGSFPQLLHLASLGILKICICRHFHSAGSLPILQPLIASI 336
QY 302 PLGSWVPKIGDINSRLMLRNPNASLSSALNMLVQAGVSSIPVDNDSDLLDYSRSDIT 361
D 337 PLGTWVPNGDPNGQLIRLRPNASLGDLMSFVQAKVSSIPVDNDSDLLDYSRSDIT 396
QY 362 ALAKDKVYTHVRIDEMTHQALQLODANTPFQFNQRCOMCLRSDDLKWKMERLANPG 421
D 397 ALAKDKAVARISLIDETNIHQALIGDANSPYGLNNGHCHNCLRSDSLKWKMERLAKG 456
QY 422 VRRVIVEAGSKRVEGIIISLSDIFKELL 449
D 457 VRRVIVEAGSKRVEGIIISLSDVFRFL 484

RESULT 6
ID C9FV59 PRELIMINARY; PRT; 382 AA.
AC C9FV59;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Putative activator subunit of SNF1-related protein kinase SNF4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20387008; PubMed=10929106;
RA Kleinow T., Shalerao R., Breuer F., Umeda M., Salchert K., Koncz C.;
RT "Functional identification of an arabidopsis snf4 ortholog by
RT screening for heterologous multicopy suppressors of snf4 deficiency in
RT yeast.";
RL Plant J. 23:115-122 (2000).
DR EMBL; AF253335; AAG10141.1; -.
DR GO; GO:0016301; P:kinase activity; IEA.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 3.
KW Kinase.
SQ SEQUENCE 382 AA; 41800 MW; A9B7A4D5E1A3CB53 CRC64;

Query Match 57.0%; Score 1334.5; DB 10; Length 382;
Best Local Similarity 58.6%; Pred. No. 1.1e-104;
Matches 262; Conservative 52; Mismatches 61; Indels 7; Gaps 4;

QY 71 SPSTPGSRMMDVDNENFQRTVTLSDGTVSEGLRVSEAAIQI-----SRC 116
D 7 SPETLG-RSNMDVD-DVFLRTADPS-----QEAVERMSGVDELRSRISVLLSTRAYEL 60
QY 131 LPDSGVKVALDINLPVKQSFHILHEQICIPVAPLWDSFRGQFVGLSPDLFILRLRELETH 190
D 61 LPESGVKVALDINLPVKQAFHILYEQICIPVAPLWDSFRGQFVGLSPDLFILRLRELETH 120
QY 191 GSNLTHEQLETHIISAWKAKQTNGRNDSQWRP--QQHLVHATPYESLRDIAVKLLQNGI 249
D 121 GSNLTHEEELIETHIAAWKEGAHISRQYDGSGRPYPRPLVQVGYDNLKVALKILQNKV 180
QY 250 STVPVIYSSSDGSPQLLHLASLGILKICICRYFNSTGNLPIINQPVCSIPGWSVVPK 309
D 181 AAVPVIYSSSDGSPQLLHLASLGILKICICRYFRHSSSSLPILQPICISPLGFWVPR 240
QY 310 IGDLSRLPAMLRPNASLSSALNMLVQAGVSSIPVDNDSDLLDYSRSDITAKDKVY 369
D 241 IGESSKXPATLRPHASLGALALVQAEVSSIPVDNDSDLLDYSRSDITAKDKAY 300
QY 370 THVRIDEMTHQALQLODANTPFQFNQRCOMCLRSDDLKWKMERLANPGVRRVIVE 429
D 301 AQIHLDMDTVHQALQLODASPPYGFIFNGQRCHMCLRSDSLKWKMERLANPGVRRVIVE 360

```

```

QY 430 AGSKRVEGIIISLSDIFKFLSL 451
D 361 AGSKRVEGIIISLSDVFRFLGL 382

```

RESULT 7

```

ID 004028 PRELIMINARY; PRT; 391 AA.
AC 004028;
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE F7G19.11 protein.
GN F7G19.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Osborne B.I., Vysotskaia V.S., Toriumi M., Yu G., Oji, O. Suehler B.,
RA Conway A.B., Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y.,
RA Shinn P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,
RA Theologis A.;
RT "The sequence of BAC F7G19 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Theologis A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; ACC00106; AAB70406.1; -.
DR PIR; B86222; B86222.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 2.
DR SMART; SM00116; CBS; 1.
SQ SEQUENCE 391 AA; 43032 MW; 26D6D08280587B74 CRC64;

Query Match 44.0%; Score 1030.5; DB 10; Length 391;
Best Local Similarity 55.0%; Pred. No. 7.4e-79;
Matches 216; Conservative 49; Mismatches 73; Indels 55; Gaps 7;

QY 71 SPSTPGSRMMDVDNENFQRTVTLSDGTVSEGLRVSEAAIQI-----SRC 116
D 7 SPETLG-RSNMDVD-DVFLRTVTS-----RVASSIWSLILIRGMLTRPRKFLPGC 53
QY 117 RVSEYLNLTCTVDLLPDSCKVIALDIN---LPVKQSFHILHEQ----- 156
D 54 RGLWSFLVTAIFYCQPALHMSCSQNRARLSVKNLMLTRKLSKVIGNFTGLLYLPIT 113
QY 157 -----GIPVAPLWDSFRGQFVGLSPDLFILRLRELETHGSLNLTAEQLETHIISAWKE 209
D 114 LYSASVGIPLAPLWDFGKQFVGLGPDLDLILRLGTHGSLNLTAEQLETHIISAWKE 173
QY 210 AKRQTNGRNDSQWRP--QQHLVHATPYESLRDIAVKLLQNGISTVPVIYSSSDGSPQLL 268
D 174 GKAHISRQYDGSGRPYPRPLVQVGYDNLKVALKILQNKVAAPVVIYSSSDGSPQLL 233
QY 269 HLASGLILKICICRYFNSTGNLPIINQPVCSIPGWSVVPKIGDINSRLPAMLRPNASLS 328
D 234 HLASGLILKICICRYFRHSSSSLPILQPICISPLGFWVPRIGESSKXPATLRPHASLG 293
QY 329 SAINMLVQAGVSSIPVDNDSDLLDYSRSDITAKDKVYTHVRIDEMTHQALQLOD 388
D 294 SALALLVQAEVSSIPVDNDSDLLDYSRSDITAKDKAYAQIHLDMDTVHQALQLOD 353
QY 389 ANTPGFFNGQRCHMCLRSDDLKWKMERLANPG 421
D 354 ASPPGYIFNGQRCHMCLRSDSLKWKMERLANPG 386

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
  Lehmann B., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
  Tümmig B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
  "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
  Nature 418:79-85(2002).
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC115579; AAC050977.1; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Hypothetical protein.
SQ SEQUENCE 577 AA; 64209 MW; 3FBC56EA649B25A9 CRC64;

Query Match 17.3%; Score 405; DB 5; Length 577;
Best Local Similarity 29.1%; Pred. No. 1.7e-25;
Matches 139; Conservative 74; Mismatches 168; Indels 96; Gaps 17;

QY 34 VDGEWRHDERQPTISGEGIVNTLYL-----TREYNQNTL-----S 70
DB 130 VDGEKGNKESQSPNGDNQILNNNNFFKDTSLPSTDNKSSTNNNNNNENPLKQTISS 189

QY 71 SPS-----TPG-----SRMMVDVNEFCETVTLSDGTVEGTLRV----- 106
DB 190 SPSKSTTTTTTTTSTTTTSPSLSSSSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 249

QY 107 -----SEMAIQISRCRVSEYLNLTCTDILDPGKVLALDINLPVKOSFHILHEQGPV 160
DB 250 KKIDSEKITEEGKQVFNELKGTCTDVIPIGKVVVLTDLKAVKSAFALENGIKS 309

QY 161 APLWDSFRGQVGLLSPDFT-LILRELETHGSKNLTQEQLTHTISA-WKAKRQTNGRN 218
DB 310 APLWNSQHDFTGMTVSDFDILLYYKPKSNNIFQDMGHRHRETETWREI----- 361

QY 219 DSQWRPOHLVHATPYESLRDIANKLQNGISTVPVIYSSSDSGSFPOLLHLASLSGLK 278
DB 362 -SVERPSS-LISTEPETNLXDAASLLCYKHLRFLPVVDKDTN-----SILHLTHSRILA 415

QY 279 CICRYEKSTGNLPIAQVCSIPLGSWVPKIGDLNSRPLAMRPNASISSALNMLVQAG 338
DB 416 FMMKSF-----POLPEKLSIFGS-----LGITGFATVTVMTHTPLVEVLELLSEKK 464

QY 339 VSSIPVDDNDS-LDITYSRSDITALAKDKVYTVRDEMTIHOAL-----QGGQDANTFF 393
DB 465 ISAVPIIDSETSKXIVDVYKSDVTLMKQKILSPDLN-LPVHVLSTFTTKLWQ----- 517

QY 394 GFFNGQRQCMCLRSDPLLKWMERLANPGRVRFVEAGSKVEGSIISLDFEKLFS 450
DB 518 ---RPEQITCTREFDKGLDVIERCIRKVRHLVLCIDS-SKKVEGLSLSDILNLYLLN 570

RESULT 10
Q8BIQ9 PRELIMINARY; PRT; 443 AA.
ID Q8BIQ9
AC Q8BIQ9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to AMP activated protein kinase gamma 1.
GN 2410051C13RIK.
OS Mus musculus (Mouse).

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
  Lehmann B., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
  Tümmig B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
  "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
  Nature 418:79-85(2002).
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC115579; AAC050977.1; -.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Hypothetical protein.
SQ SEQUENCE 577 AA; 64209 MW; 3FBC56EA649B25A9 CRC64;

Query Match 35.0%; Score 820.5; DB 10; Length 451;
Best Local Similarity 39.5%; Pred. No. 6.2e-61;
Matches 177; Conservative 86; Mismatches 150; Indels 35; Gaps 6;

QY 13 CP-----TVFQAICSLSPGHEHYKFFVDGGEWRHDERQPTISGEGIVNTLYLTREYNQIN 67
DB 27 CPWGLGAEQVVDLPFGVYQVREFLVGVWRCDKTPCVRDEYGLISNEVLVDNTHPV- 85

QY 68 TLLSPSTPGSRMMMDVNEFNQRTVTLSDGTV-----SEGLTVSEAAIQISRCRVSEYL 122
DB 86 -----VQPEITSIRVSMDEGTLITMPEPDQLSQNSGVQIAIFRERVSEIL 130

QY 123 NLHTCYDLLPDSGVKVIADNLNPKOSFHLIHEQGIIPVAPLWDSEFRQGVGLLSPDLFIL 182
DB 131 LEMTIYDVVPSSKIAVLDAKLPVKQAFKIMHDEGLSIVPLWDQDQQTGTGMLTASDFVL 190

QY 183 ILRELETHGSKNLTQEQLTHTISA-WKAKRQTNGRNDSQWRPOHLVHATPYESLRDIAP 242
DB 191 ILAKLQNRNIRTLGHELEHMSVSAWKAQLQFYCGPDVAALQRRPLIHVKDSNLRDVAL 250

QY 243 KLLONGISTVPVIYSSSDSGSFPOLLHLASLSGLKICRYFNKSTGNLPIAQVCSIP 302
DB 251 AITRNEISSVP-IPKPSDSSGMPLGLATLPGVRFKICSLQEQPBGYSFLQNGIVSMP 309

QY 303 LGSWVPKIGDLNSRPLAMRPNASISSALNMLVQAGVSSIPVDDNDSLDTYSRSDITA 362
DB 310 IGWSPHTGKASNRQLRTSRPSTPLNSCLLLEDRVSSIPVDDNGLLDVYLSOIMA 369

QY 363 LAKDKVYTVRDEMTIHOALQGDQATPPGFNGQR-COMCLRSDPLLKWMERLANP 421
DB 370 LGRNDVYTRIELEQVTVHEALELQYQV-----NGRRHCHTCLSTSTFLEVLQLSAPG 422

QY 422 VRRVFIAGSKRVEGSIISLDFEKL 449
DB 423 VRRVVVTEPRSRFVQGIISLRDAFTFLI 450

RESULT 9
Q86KN7 PRELIMINARY; PRT; 577 AA.
ID Q86KN7
AC Q86KN7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

```

OC Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT the RIKEN Genome Expression Transcriptome Research Group Phase I & II Team;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR ENBL; AX032238; BAC27775.1; -.
DR MGD; MGI:1920953; 2410051C13Rik.
DR InterPro; IPR000644; CBS domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SMC0116; CBS; 4.
SQ SEQUENCE 443 AA; 49980 MW; AC336AD3FE35E28 CRC64;

Query Match 16.1%; Score 376.5; DB 11; Length 443;
Best Local Similarity 28.0%; Pred. No. 2.9e-23;
Matches 117; Conservative 79; Mismatches 143; Indels 79; Gaps 17;

Qy 61 REYNQINTLSSPSPGSRMMNDVNRNQ-----RTVTL--SDGTVSEG 102
Db 62 RPESRIYASSPPTGQRFCL----AFQSPARPPLASETHAPLRTAVLAAPGPAEAG 116

Qy 103 TLRV-----SEAAIQISRCRVSEYLNHTCYDLLPDSGKVIALDINLPVKSHEIL 153
Db 117 MLEKLEFQEEADESGGYM-----RFRSHKCYDIVPTSSKLVVPTTIOVKKAFAL 170

Qy 154 HEQIGVPAPLWDSPRGQFVGLLSPLDFILIRELTHETSGNLTE--EQLETHTISAKFAKR 212
Db 171 VANGVRAAPLWESKKQSFVGLMTITDFINILHRY--YKSPMVQIYELEEHKIEWRELYL 228

Qy 213 QTNGRNDSQWRPQOHVHATPYESLRDIKVLKLONGISTVPIVYSSSDSGSPQLLHLAS 272
Db 229 Q-----ETFKP---LVNISPDASLDFQAVYSLIKNKHRLFPVIDPISGNA-----LYILT 274

Qy 273 LSGILKICIRYKFNKSTGNLPILNQPVCISIPIGSMVPKIGLNSRPLAMLRPNASLSSALN 332
Db 275 HKRIKFL-QLFMGDMKPAFMKNIDELGLTV-----HNIAFIHPTPIIKALN 324

Qy 333 MLVOAGVSSIPDIVDNDLSLDTYSRSDITALAKDKVYTHVELDMEITHQALQLOGDANTP 392
Db 325 IFVERRISALPVVDSESGKVDIYKSFQDVINLAAEKTYNNL---DITVTCALQHRSQ----- 377

Qy 393 GGFENGQRQCMCLSDPLLKWMERLANPGVRVFIV--EAGSKRVEGIISLSDIPKFL 449
Db 378 --YFEG--VVKCKSLKETLTIIVDRIVRAEVRHLVVVNEADS--IVGIISLSDILQALI 429

RESULT 11
Q91WGS PRELIMINARY; PRT; 566 AA.
ID Q91WGS
AC Q91WGS
DT 01-DEC-2001 (TrEMBLrel. 19, Created);
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update);
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update);
DE Hypothetical protein.
GN 2410051C13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA Strausberg R.;
DR ENBL; BC015283; AAH15283.1; -.
DR MGD; MGI:1920950; 2410051C13Rik.
DR InterPro; IPR000644; CBS domain.
DR

Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kemrison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy W., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusser D.K., Pacle C.M.,
Palazzolo M., Pittman G.S., Pat S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.Y., Wasarman D.A., Weinstein G.M., Weissensbach J.,
Williams S.M., Woodager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of *Drosophila melanogaster*.";
Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RP
CElniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
Evans C.A., Gocayne S.E., Amaratides P.G., Brandon R.C., Rogers Y.,
Banzon J., An H., Baldwin D., Banson J., Beeson K.Y., Busam D.A.,
Carlson J.W., Center A., Champé X., Davenport L.B., Dietz S.M.,
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
Ferreira S., Frise E., Gallie R.F., Garg N.S., George R.A.,
Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
"Sequencing of *Drosophila melanogaster* genome.";
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
[3]
SEQUENCE FROM N.A.
RP
Misra S., Crohey M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
Aratecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnikner S.E.,
Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
"Annotation of *Drosophila melanogaster* genome.";
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
[4]
SEQUENCE FROM N.A.
RP
Adams M.D., Celnikner S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
[5]
SEQUENCE FROM N.A.
RP
FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR ENBL; AS003733; AN13855.1; -
DR FlyBase; FBgn0025803; SNF4A-gamma.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
SQ SEQUENCE 585 AA: 64723 MW: 78FF8F67E841EAF CRC64;

Query Match 15.8%; Score 369; DB 5; Length: 585;
Best Local Similarity 27.0%; Pred. No. 1.9e-22;
Matches 120; Conservative 85; Mismatches 153; Indels 86; Gaps 16;

CY 45 PTISGEFGIVNTLYL-----TREYNQWTLS-SPSTPGSRMMVDVNFQ 89
DB 6 PTASSNYQLNSANLRDLIDQRHSHVRRDRDSQSHSESFSHG---HHSQNQSIP 62
CY 90 RTVTYL-----SDGTVSGETLVRSAAIQISCRVSV-----EYLNHTCTDCLL 131

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Meinkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.C., Nusser D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassaman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodgeri, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*";
RL Science 287:2185-2195(2000).
RP SEQUENCE FROM N.A.
RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champs M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Gallie R.F., Gary N.S., George R.A.,
RA Gonzalez X., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bertram B., Carlson J.W., Celnik S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY084138; AL89876.1; -
DR EMBL: AE003733; AN1382.1; -
DR FlyBase: FBgn0025603; SNF4A-gamma.
DR InterPro: IPR000644; CBS_domain.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
SQ SEQUENCE 906 AA; 99970 MW; C867D9556F42D57F CRC64;
Query Match 15.6%; Score 366; DB 5; Length 906;
Best Local Similarity 26.5%; Pred. No. 6.8e-22;
Matches 122; Conservative 81; Mismatches 164; Indels 94; Gaps 17;
QY 39 RHDEROPTISGEFGIVNTLYLTREYN-QINTLSPPS-----TPGSRM- 79
Db 307 KKHDAQSVCTSPQNDQENFRTRKSLVDNLTSLRPNILDDCGATKYIVRIKGAHMR 366
QY 80 NMDVDNENFQRTV-TLSDGTIVSEGTIVRVEAATQ-----ISCRVSE----- 120
Db 80 NMDVDNENFQRTV-TLSDGTIVSEGTIVRVEAATQ-----ISCRVSE----- 120

Db 367 RASADLEKRRASVGAAGRGLRGDTGDTHDHAAILFRDSRGLPVAOPFLEKVNLSLEED 426
QY 121 -----YLNLTCTYLLDSDSGKVIADINLPVKQSFHILHEOGIPVAPLWDSFRGQFVGL 174
Db 427 SQIFVKFFREHKCYDLIPTSAKLVVDFDQLLVKAFYALVYGVRAAPLWDSKQOFVGM 486
QY 175 LSPDLFILILRELETHSGNLTBEQLETHITSAAKEAKQRTNGRNDSSQMPQOHLVHATFY 234
Db 487 LTITDFIKL-QMYKSPNASNEQLEHKLDTWRSVLH-----NQVMP---LVSIGPD 535
QY 235 ESLRDIKVLQNGISTVPVIYSSSDSGSPFQOLLHLASGLKICICRYFKNSTGNLPIL 294
Db 536 ASLYDAIKLILHSRIHPLVI-----DPATGNVLYILTKRILRFLFY----- 579
QY 295 NQPVCSIPGWSVP-----KIGDLSNRPLAMRPNASLSSALNMLVQAGVSSIPVDN 348
Db 580 ----INELPKPATYQKSLREKLTGTNN--IFTADETSIITALKKEVERVSALPLVDS 634
QY 349 DSLSDTYSRSDITALAKOKVYTHVRIDEMTHIQALQGLQDANTPFGFFNGQRCQMLRSD 408
Db 635 GRLDVIYAKEDVINLAAEKTNDL---DVSRLKANEHRNE-----WFBG--VQKCNLDE 683
QY 409 PILKWERLANPGVRVRFIVEAGSKRVEGIIISLDFIKELL 449
Db 684 SLYTIMERIVRAEVHRLVVVDE-NRKVIGIISLSDILLYLV 723

RESULT 14
Q9BHL6 PRELIMINARY; PRT; 582 AA.
AC Q9BHL6; DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Y11B2A.8.
OS Y11B2A.8.
GN *Caenorhabditis elegans*.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; *Caenorhabditis*.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RX none;
RA "Genome sequence of the nematode *C. elegans*: A platform for
RT investigating biology.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RC Sulston J.E.;
RA Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL132904; CAC35836.2; -
RW Hypothetical protein.
SQ SEQUENCE 582 AA; 64153 MW; D0FC77F406EA2C6 CRC64;
Query Match 15.6%; Score 365; DB 5; Length 582;
Best Local Similarity 28.6%; Pred. No. 4.2e-22;
Matches 104; Conservative 73; Mismatches 135; Indels 52; Gaps 14;
QY 94 LSDGTIVSEGTIVRVEA-----AIOISRCRVSEYINLHTCYDILLPDSKVIADINLPVKQS 149
Db 185 LSDPNFDFTWHTSEADNHDAV-----YSLFMAHKCYDLIPTSKLVVFTHLFVRKA 238
QY 150 PHILHGOQFVAPLWDSFRGQFVGLASPLDILILRELETHSGNLT-EQLETHITSAAWK 208
Db 239 FYALVYGVRAAPLWDTDNRFTGMLTITDIFKILCKHYDKGNSEIRALEQQISHWR 298
QY 209 EAKRQTNGRNDSQWRPQOHLVHATPYESLRDIKVLQNGISTVPVIYSSSDSGSPFQOLL 268
Db 299 D-----QFELDGTLRP---FVIDPNESLHRAVELLCSKVHPLVLD-----DRKTGNIT 345

```
QY 269 HLASLSGILKACICRYFKNSTGNLPIINOPVCS---IFLGSWVPKIGDLSRPLAMLRPN 325
DB 346 YILTHKIMKFLSYMR-----DLPRFMSCTPRELIGAW---GD-----ILCHVDVT 392
QY 326 SLSSALNMLVQAGVSSIPVDDNDLSLETYSRSDITALAKDKVYTHVRLDENTHQAQL 385
DB 393 PIHDALEFLKNRVSLPIDENGRVVDIYAKFDVISLAESSYDKL---DCTVQEAHQH 449
QY 386 QDANTPFPGFNGQRCQCLRSDELLKMEELANPGVRRVPIVEAGSKRVEGIIISLSDIF 445
DB 450 RSE-----WFEG--VQTCLETSLFQVLEAIVAEVHRLLIVTQC-DKKVGVVVSLSL 500
QY 446 KELL 449
DB 501 KLV 504

RESULT 15
O96613 PRELIMINARY; ERT; 647 AA.
AC O96613
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SNF4/AMP-activated protein kinase gamma subunit.
GN SNF4A-GAMMA OR CG5806 OR CG17299.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Yoshida E.N., Benkel B.F., Fong Y., Hickey D.A.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF094764; AAC95306.1; -.
DR EMBL; AF094763; AAC95305.1; -.
DR FlyBase; FBgn0025803; SNF4A-gamma.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Kinase.
SQ SEQUENCE 647 AA; 71592 MW; B792BE1089730B52 CRC64;

Query Match 15.5%; Score 363.5; DB 5; Length 647;
Best Local Similarity 28.2%; Pred. No. 6.6e-22;
Matches 108; Conservative 72; Mismatches 123; Indels 74; Gaps 13;

QY 97 GTVSGEGLRVGEAAIQ-----ISGRVSE-----YLNHFTCYLLP 132
DB 126 GLRGDGLDPHAAILFEDSRGLPVADPFLEKVNLSDEEDDSQIFVKAFRFHKCYLLP 185
QY 133 DSGKVIALDINLPVKQSHILHEQIPVAPWDSFRGQFVGLLSPDLPILIRLEIETHGS 192
DB 186 TSAKLVPFDTQLLVKAFALVYNGVRAAPLWDSEKQFVGMLTITDFIKIL-QMYKSP 244
QY 193 NLTEBQLECHTISAWKEAKRQTNGRNSQWEPQOHLVHATPYESLRDIKVLQNGISTV 252
DB 245 NASMEQLEEHKLDTRDVIH-----NQVMP---LVSIGEDASLYDAIKILHSRIHRL 294
QY 253 PVIYSSSSSDGSPFOLLHLASLSGILKICRYFKNSTGNLPIINOPVCSIPLGSWVP--- 308
DB 295 PVI-----DPATGNVYIILTHKRIILFLFLY-----INELPKPAYMKSIR 335
QY 309 --KIGDLSRPLAMLRPNASLSALNMLVQAGVSSIPVDDNDLSLDTYSRSDITALAKD 366
DB 336 ELKIGTYNN--IETADETTTSIITALKKFRVRSVSLPLVDSGRLVDIYAKFDVINLAAE 393
QY 367 KYTHVRLDENTHQAQLQGDANTPGCFNGQRCQCLRSDELLKMEELANPGVRRV 426
DB 394 KTYNDL---DVSRLKANEHRNE-----WFEG--VQCNLDESLYTIMERIVRAEVHRLV 442
```

```
QY 427 IVEAGSKRVEGIIISLSDIFKELL 449
DB 443 VVDE-NKXVIGIISLSDILLYLV 464

Search completed: July 7, 2004, 17:52:19
Job time : 48 secs
```